JOURNAL OF APPLIED PHYSICS

JOURNAL
OF
APPLIED PHYSICS

Vol. 23, No. 10 OCTOBER, 1952

In This Issue

Page

Experimental Study of the Formation of a Vortex Ring at the Open End of a Cylindrical Shock Tube F. K. Elder, Jr., and N. de Haas 1065

Property of Bessel Functions and Its Application to the Theory of Two Rheometers Hershel Markovitz 1070

Concept and Measurement of Impedance in Periodically Loaded Wave Guides Edwin T. Jaynes 1077

Complete Solution of the One-Velocity Diffusion Problem at Intermediate and Large Distances Lewi Tonks 1085

Condensation of Nitrogen in a Hypersonic Nozzle W. W. Willmarth and H. T. Nagamatsu 1089

Announcement 1095

Influence of Initial Velocities on Electron Transit Time in Diodes J. T. Wallmark 1096

On a Multiple Scattering Theory of the Finite Grating and the Wood Anomalies V. Twersky 1099

On the Use of the Joule-Thomson Effect to Measure the Work Equivalent of Heat. Thermodynamic Theory John R. Roebuck 1119

On the Electrical Properties of Porous Semiconductors Eugene B. Hensley 1122

Electrical Conductivity and Thermoelectric Power of (BaSr)O and BaO J. R. Young 1129

Method of Solving the Wave Equation in a Region of Rapidly Varying Complex Refractive Index J. J. Gibbons and R. L. Schrag 1139

Announcement 1142

Note on the Approach of Narrow Band Noise after a Nonlinear Device to a Normal Probability Density George R. Arthur 1143

Electronic Properties of Barium Sulfide Walter Gratthidge and Harold John 1145

Scattering of Electromagnetic Energy in a Randomly Inhomogeneous Atmosphere Harold Staras 1152

Microchronometric Schlieren Study of Gaseous Expansion from an Electric Spark H. L. Olsen, R. B. Edmonson, and E. L. Gayhart 1157

Effect of Internal Strains on Linear Expansion, X-Ray Lattice Constant, and Density of Crystals P. H. Miller, Jr., and B. R. Russell 1163

Bonding of Thin Films H. P. Meissner and J. Byrne 1170

Theory of Electrically Short Transmitting and Receiving Antennas Ronold King 1174

Letters to the Editor:

Phase Changes in Silver—Tin Amalgams J. C. Moffett, G. Ryge, and A. G. Barkow 1188

Effect of Orientation on the Ultimate Strength of Linear High Polymers C. C. Hsiao 1189

On the Oscillating Cylinder Viscometer. I Ali A. K. Ibrahim and A. M. Kabiel 1190

Index to Advertisers xxxiv

BOARD OF EDITORS
Elmer Hutchisson, *Editor*
Earle C. Gregg, *Assistant Editor*

Associate Editors
Howard H. Aiken
Ray E. Bolz
Lyle D. Borst
Robert D. Heidenreich
James Hillier
Leonard B. Loeb
W. James Lyons
Eric Reissner
Roman Smoluchowski

Publication of the
American Institute of Physics

OFFICERS
George R. Harrison, *Chairman*
George B. Pegram, *Treasurer*

GOVERNING BOARD
Karl K. Darrow
G. J. Dienes
S. A. Goudsmit
G. P. Harnwell
George R. Harrison
F. V. Hunt
Rudolph Kingslake
Hugh S. Knowles
C. C. Lauritsen
F. W. Loomis
William F. Meggers
George B. Pegram
I. I. Rabi
Duane Roller
Frederick Seitz
Francis G. Slack
John C. Steinberg
Richard M. Sutton
Mark W. Zemansky

Consultant on Printing
Melvin Loos

ADMINISTRATIVE STAFF
Henry A. Barton, *Director*
Wallace Waterfall, *Executive Secretary*
Theodore Vorburger, *Advertising Manager*
Ruth F. Bryans, *Publication Manager*
Betty H. Goodfriend, *Assistant Secretary*
Eileen B. Neuberger, *Circulation Manager*
Kathryn Setze *Assistant Treasurer*

The Journal of Applied Physics, published monthly at Prince and Lemon Sts., Lancaster, Pa., is devoted to physics in its role as the science basic to other natural sciences and to the arts and industries. Previous to 1937 this publication was known as Physics. The Journal of Applied Physics publishes editorials and reviews, as well as technical papers of applied physics. Articles appearing in it are indexed regularly, according to their field, by Physics Abstracts, Chemical Abstracts, Engineering Index, Inc., Industrial Arts Index, etc.

Manuscripts should be submitted to Elmer Hutchisson, Editor, Case Institute of Technology, Cleveland 6, Ohio. Unless otherwise stated, submission of a manuscript is a representation that it has been neither copyrighted, published, nor submitted for publication elsewhere. The authors' institutions are requested to pay a **publication charge** of **$8** per page which, if honored, entitles them to 100 reprints without further charge. Instructions will be sent with galley proofs.

Subscriptions, renewals, and orders for back numbers should be addressed to the American Institute of Physics, 57 East 55 Street, New York 22, New York.

Subscription Price	U.S. and Canada	Elsewhere
To members of the American Institute of Physics and Affiliated Societies	$10.00	$11.00
To all others	12.00	13.00

Back Number Prices

Yearly back number rate when complete year is available: **$14.00.** Single copies: **$1.25** each.

Changes of address should be sent to the American Institute of Physics; **corrected proofs** should be sent to the Publication Manager, American Institute of Physics. **Advertising rates** supplied on request. Orders, advertising copy, and cuts should be sent to the American Institute of Physics.

Entered as second class matter January 22, 1937, at the Post Office at Lancaster, Pennsylvania, under the Act of March 3, 1879. Accepted for mailing at the special rate of postage provided for in paragraph (d-2), section 34.40, P. L. & R. of 1948, authorized May 2, 1932.

Journal
of
Applied Physics

Volume 23, Number 10 October, 1952

Experimental Study of the Formation of a Vortex Ring at the Open End of a Cylindrical Shock Tube*

F. K. Elder, Jr., and N. de Haas
Applied Physics Laboratory, The Johns Hopkins University, Silver Spring, Maryland
(Received March 5, 1952)

Schilieren spark snapshots spaced at electronically controlled intervals of several hundredths of a millisecond show the emergence of the shock wave and the formation of the vortex ring at the end of a cylindrical shock tube, open to the atmosphere. The axial positions of the vortex ring and shock front are plotted as a function of time for the first millisecond, in which time the vortex forms and is accelerated to about three-quarters of the theoretical particle velocity of the mass flow following the initial plane shock. The diameter of the vortex ring increases nonlinearly with time and distance.

INTRODUCTION

THE study of vortex rings lost much of its glamor for physicists with the passing of Thomson's vortex atoms.[1] The "steady-state" behavior of a vortex ring moving with nearly constant velocity in a fluid has been treated theoretically in classical treatises such as Lamb's *Hydrodynamics*,[2] as well as in the periodical literature. The generation of a vortex ring by an impulsive pressure acting over a circular area has been discussed by von Kármán and Burgers,[3] who derive an expression for the strength of such a ring. However, there appears to have been no previous study, either experimental or theoretical, of the transient behavior of a vortex ring in air during its formation and initial acceleration.[4] This paper presents the results of an experimental observation of the formation and acceleration of a vortex ring at the open end of a cylindrical shock tube, during the first millisecond after emergence of the shock front.

Gawthrop, Shepherd, and Perrott[5] in 1931 established the existence of the vortex ring set up at the orifice of a tube by an emerging shock wave, the shock wave having been produced by the detonation of an explosive charge at the opposite end of the tube. Payman and Shepherd,[6] investigating in some detail the properties of the shock tube, have published some interesting schlieren photographs of the structure of the pressure effects projected from the end of the tube into free air, at compression-chamber pressures of 430 and 120 psi. No attempt was made in either of these articles to study in detail the formation and initial acceleration of the vortex ring. A study of the formation and propagation of vortex rings in air was reported in 1926 by Sadron,[7] but his mass-flow velocities were less than ours by a factor of the order of 100, and his time measurements were taken in intervals of the order of tenths of a second. With the benefit of electronic timing we have been able to make a number of observations at controlled intervals in the first millisecond after emergence of the shock wave. We are thus concerned with the initial acceleration of a vortex ring formed in air behind an expanding shock, whereas Sadron was

* This research was supported by the Bureau of Ordnance, U. S. Navy, under Contract NOrd-7386.
[1] J. J. Thomson, *On the Motion of Vortex Rings* (Adams Prize Essay, London, 1883).
[2] H. Lamb, *Hydrodynamics* (Cambridge University Press, Cambridge, 1932); sixth edition.
[3] T. von Kármán and Burgers, *Aerodynamic Theory*, W. F. Durand, editor (1935), Vol. II.
[4] Of interest in this connection, however, are various observations of diffraction of shock waves around rectangular obstacles in two-dimensional flow in a shock tube. The development of line vortices in the course of this diffraction is shown rather nicely, for example, in such papers as Bleakney, White, and Griffith, J. Appl. Mech. **17**, 439 (1950); Hollyer and Hunting, University of Michigan, Engineering Research Institute Report 51–4, ONR Project M720–4 (1951).
[5] Gawthrop, Shepherd, and Perrott, J. Franklin Inst. **211**, 67 (1931).
[6] Payman and Shepherd, Proc. Roy. Soc. (London) **A186**, 293 (1946).
[7] Sadron, J. phys. **7**, 76 (1926).

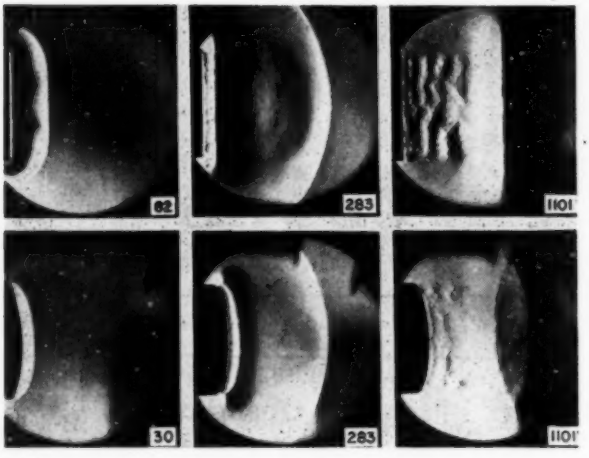


FIG. 1. Schlieren snapshots of shock front and vortex ring at open end of cylindrical shock tube (40 psi compression). Numbers indicate time after shock-front emergence (microseconds). (a) Schlieren axis normal to tube axis. (b) Schlieren axis tilted at 15° to normal to tube axis.

concerned with the subsequent deceleration of a vortex ring moving through still air.

Banerji and Barave,[8] and later Krutzsch,[9] have observed a similar phenomenon in liquids—the formation of a vortex ring by the ejection of liquid from a submerged tube. By a suitable coloring method in each case, these authors have been able to trace out the streamlines in the liquid as the vortex is formed and begins to move away from the orifice. By simultaneously lightly coloring the entire volume of liquid initially within the submerged tube, Banerji and Barave were able to show not only the streamlines, but also the position of the vortex core in relation to the outer boundary of liquid initially within the tube.

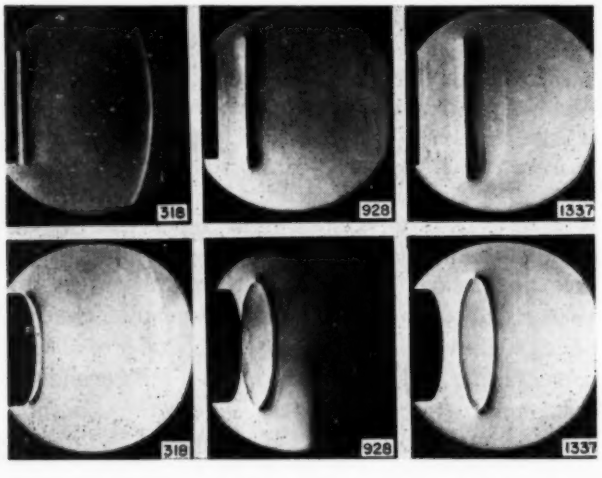


FIG. 2. Schlieren snapshots of shock front and vortex ring at open end of cylindrical shock tube (10 psi compression). (a) Schlieren axis normal to tube axis. (b) Schlieren axis tilted at 15° to normal to tube axis.

[8] Banerji and Barave, Phil. Mag. **11**, 1057 (1931).
[9] Krutzsch, Ann. Physik **35**, 497 (1939).

APPARATUS

The shock tube was made of steel pipe, 4.00 inches in outer diameter and 3.375 inches in inner diameter. The compression chamber was 24 inches long, and the expansion chamber 176 inches long. A spring-loaded needle, released manually, punctured the diaphragm when the desired pressure was reached. Diaphragms were made of one or more thicknesses of red zip tape, the number of thicknesses in each instance being the minimum which would sustain the desired pressure without spontaneously rupturing. Super-dry tank nitrogen was used to fill the compression chamber, the expansion chamber being left open to the atmosphere. The tube was flushed with compressed air after each shot to remove the pieces of the burst diaphragm.

TABLE I. Initial motion and growth of vortex ring.

10 psi compression				40 psi compression			
Time[a] (μsec)	Position[b] of core (inches)		Diameter of core (inches)	Time[a] (μsec)	Position[b] of core (inches)		Diameter of core (inches)
		15°				15°	
174	0.07			55	0.04		
225	0.10			82	0.10		
276	0.15			106	0.12		
318	0.18			130	0.16		
327	0.20			181	0.25	0.27	3.70
378	0.24			232	0.43		
417	0.30			283	0.54	0.51	4.13
429	0.31			334	0.70		
480	0.38			386	0.82	0.78	4.38
519	0.42	0.45	4.03	427	1.04		
531	0.47			488	1.20	1.17	4.56
583	0.53			539	1.43		
622	0.59			590	1.61	1.58	4.67
658		0.67	4.21	692	1.94	2.03	4.89
685	0.65			794	2.30		
724	0.73	0.72	4.24	897	2.86	[illegible]	
787	0.82			1101	3.82	3.74	5.07
863		0.95	4.30				
928	1.05	1.07	4.44				
1067		1.31	4.58				
1133	1.37	1.40	4.62				
1337	1.72	1.75	4.79				

[a] Time after emergence of shock front from shock tube. Owing to the uncertainty in establishing this instant of emergence, there is a possible error of the order of 2 μsec. The error in time differences, however, should be less than 1 μsec.
[b] Distance from the end of the shock tube. First column is measured from normal photographs, second column measured from 15° photographs used to obtain diameters.

The schlieren system made use of a pair of 8-inch parabolic mirrors, of focal length 80 inches. The knife-edge and slit were vertical. A spark source of about a microsecond duration supplied the illumination, a 4×5 view camera being used to take the pictures.

Timing was accomplished with the use of an electronic time-delay circuit[10] adjustable to the nearest microsecond, which initiated the illuminating spark for the photograph. The delay circuit was triggered with an electric pulse from a pressure switch, mounted at the end of a vertical one-quarter-inch i.d. tube which tapped the shock tube 43 inches from the open end. A pressure pulse generated by the shock wave passing this point in the shock tube actuated the pressure switch.

[10] See Edmonson, Gayhart, and Olsen, Phot. Eng. **3**, No. 3 (1952).

EXPERIMENTAL RESULTS

As the shock tube was of a conventional type, a reasonably accurate estimate of the strength of the plane shock within the tube can be derived from the initial values of the compression-chamber and expansion-chamber pressures.[11] From this one can also estimate the theoretical mass-flow velocity behind the shock in the tube, which is approximately the initial flow velocity of the air jet[12] out the open end of the tube. Initial conditions at the end of the shock tube were investigated for two values of compression-chamber (gauge) pressure: 10 psi and 40 psi. These produce shock velocities of approximately 1250 and 1480 ft/sec, respectively. The corresponding mass-flow velocities are about 210 and 525 ft/sec. Within the limits of experimental error, the observed shock velocities are in agreement with these predicted values.

A brief survey of the first millisecond after emergence of the shock front is given for 40 psi in Fig. 1 and for 10 psi in Fig. 2. The numbers on the individual snapshots represent the time delay in microseconds (see Table I) measured from the instant of emergence of the shock wave. Each snapshot represents a separate firing of the shock tube. It should be noted that these particular snapshots do not correspond to transition points in the observed phenomena but are merely representative of successive phases. Sequence (a) was photographed normal to the tube axis, and sequence (b) is a corresponding set photographed at 15° to the normal in order to give a three-dimensional effect. At 30 microseconds (40-psi compression, Fig. 1) the 15° picture shows the shock front emerging from the end of the tube. At 82 microseconds (normal picture) the vortex ring developing at emergence of the shock front is already clearly in evidence. At 283 microseconds the shock front has nearly left the field of view, while the vortex ring is beginning to detach from the tube. At 1101 microseconds, the detached vortex ring appears at the head of a well-defined turbulent jet. At the lower pressure (10-psi compression, Fig. 2) it is observed that the vortex ring travels much more slowly. That the ring exists from the instant the shock emerges, even at this lower pressure, is evident in Fig. 3, whose first three shots precede in direct sequence those of Fig. 2(a).

Conditions Ahead of Vortex Ring

Although the shock front begins immediate three-dimensional expansion upon emergence from the tube, the axial portion of the front remains effectively plane for a distance approximately equal to the tube diameter, or until the rarefaction from the edge of the tube has caught up with the center of the shock front. The progressive effect of this rarefaction is shown in Fig. 3 for 10-psi compression. At twelve inches from the end of the tube (811 microseconds, Fig. 3) the front has become essentially spherical, with a radius of curvature of about thirteen inches. The axial portion of the front retains its initial velocity (*viz.*, 1480 ft/sec for 40 psi compression and 1250 ft/sec for 10-psi compression) until reached by the rarefaction, when it rapidly decelerates to a constant velocity very slightly greater than the speed of sound. The relative positions of the vortex ring and the axial portion of the shock front at successive instants are shown in Fig. 4.

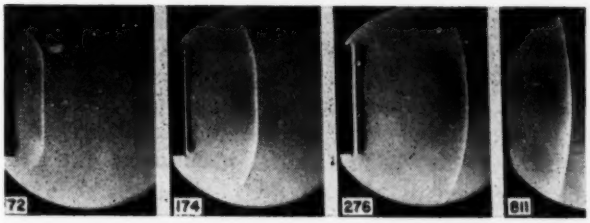


FIG. 3. Initial development of vortex ring and spherical shock front (10 psi compression). Shock front in fourth picture is 12.1 inches from end of tube.

Acceleration of Vortex Ring

The initial motion of the core of the vortex ring is plotted in Fig. 4. The ring appears to form with neg-

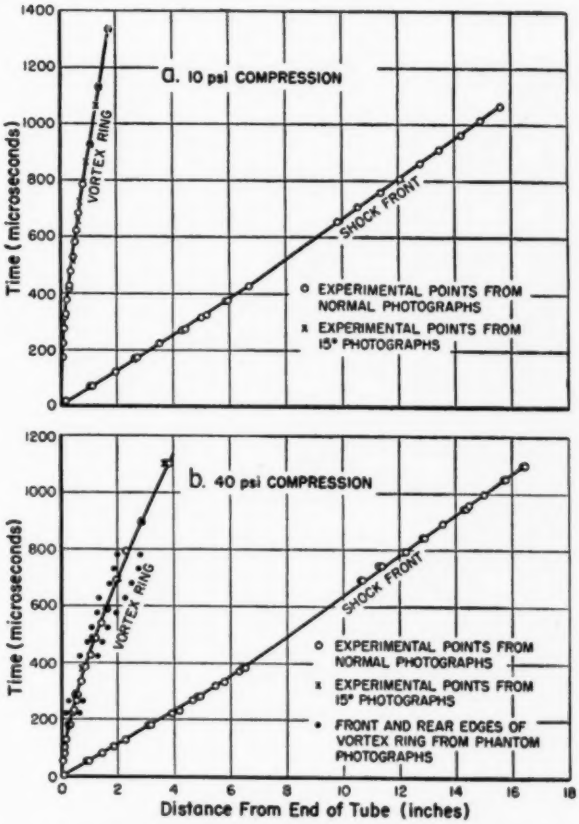

a. 10 psi COMPRESSION
VORTEX RING
SHOCK FRONT
o EXPERIMENTAL POINTS FROM NORMAL PHOTOGRAPHS
x EXPERIMENTAL POINTS FROM 15° PHOTOGRAPHS
b. 40 psi COMPRESSION
VORTEX RING
SHOCK FRONT
o EXPERIMENTAL POINTS FROM NORMAL PHOTOGRAPHS
x EXPERIMENTAL POINTS FROM 15° PHOTOGRAPHS
• FRONT AND REAR EDGES OF VORTEX RING FROM PHANTOM PHOTOGRAPHS
Time (microseconds)
Distance From End of Tube (inches)

FIG. 4. Initial motion of shock front and vortex ring at open end of cylindrical shock tube.

[11] See, for example, Bleakney, Weimer, and Fletcher, Rev. Sci. Instr. **20**, 807 (1949).

[12] In our experiments there is no nitrogen jet, but only the jet of air from the expansion chamber. The ratio of expansion-chamber length to compression-chamber length is sufficiently large that at our operating pressures the nitrogen-air contact surface never reaches the open end of the tube.

ligible axial velocity at the instant the shock front emerges from the tube and undergoes continuous axial acceleration as it increases in size and detaches from the tube. After detachment, it undergoes acceleration at a diminishing rate, approaching the theoretical particle velocity of the initial mass flow of the air emerging from the shock tube, but not reaching it in the first millisecond of its life. At 1000 microseconds the velocity is still increasing slightly, having reached about 385 ft/sec in the 40-psi case and about 135 ft/sec in the 10-psi case. The original position-time data of Fig. 4 are summarized in Table I.

Identification of Vortex-Ring Components

The appearance of the vortex ring in schlieren photographs depends very strikingly on the sensitivity of the schlieren system, as shown in Fig. 5. With the exception of (b), all the photographs in this figure were taken at essentially the same phase of the development of the vortex ring, and hence portray nearly identical conditions. The snapshots are arranged in order of increasing sensitivity. At extremely low sensitivity (a), one obtains a "phantom" photograph showing clearly the basic structure of the vortex ring. At medium and high sensitivity (c), (d), the ring appears as a very dark region followed immediately by a very bright region—as observed by previous investigators.[5,6] The rather sharp boundary between the dark and bright regions occurs at the core of the vortex. At slightly higher sensitivity than (a), the transition occurs between the phantom and ordinary schlieren vortex photographs: The rather faint dark region in (b) is to be identified with the more intense dark region observed at higher sensitivity, but there is at the lower trailing edge a circular arc suggestive of the phantom ring cross section. Comparison of (c) and (d), which were taken at the same time-delay, indicates that the apparent outer boundary of the vortex ring is a function of schlieren sensitivity.

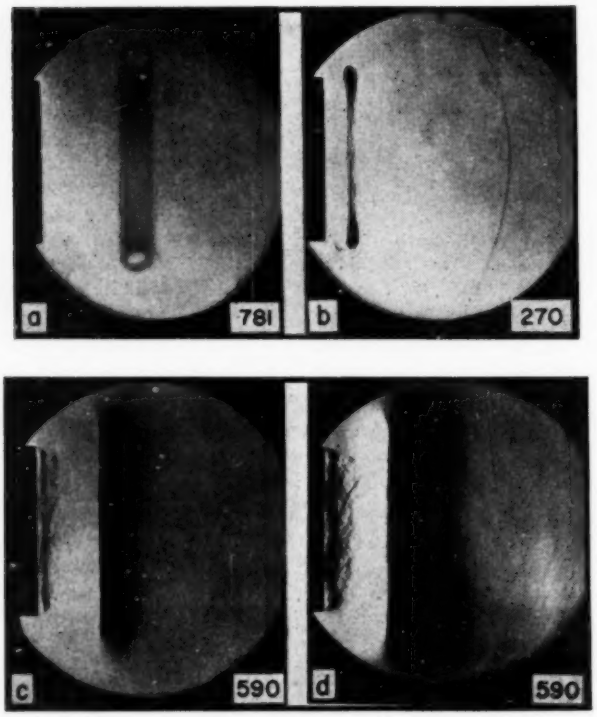


FIG. 5. Effect of schlieren sensitivity on observed details of vortex ring (40-psi compression). Photographs are in order of increasing sensitivity. The first two pictures, (a) and (b), have been retouched to offset loss of detail in reproduction.

In the 15° photographs (Figs. 1(b) and 2(b)) the bright region does not appear as clearly as the dark region, but the core shows up clearly as the rear inner boundary of the dark ring. This identification of the core is confirmed by the fact that the position of the vortex core as determined from the 15° photographs agrees well with the results from the normal photographs as shown in Fig. 4(b). Also plotted in Fig. 4(b) are measurements of the positions of the front and rear edges of the basic structure of the ring at various instants, obtained from the phantom photographs.

Size of Vortex Ring

The diameter of the vortex core as a function of time is given in Fig. 6. These measurements, recorded in Table I, were derived from the 15° pictures. In the 40-psi case, the inner and outer diameter of the vortex ring as determined from the phantom snapshots are plotted for comparison. It is evident that the diameter of the ring is a nonlinear function of distance as well as time.

An interesting phenomenon occurs in the 1337-microsecond shot in Fig. 2(a), the evidence of a mushroom-shaped head surrounding the vortex ring. Comparison with photographs of the formation of vortex rings in liquids[13] suggests that this represents the outer boundary of the air previously within the shock tube. Furthermore, the axial position of its leading edge is within about three percent of the calculated position of a particle originally just inside the end of the tube, moving with the theoretical mass velocity after the shock. This head is not evident in all of the schlieren photographs, but appears sufficiently often that it certainly exists, although its visibility must depend rather critically on test conditions.

DISCUSSION

There appears to be no simple theory directly applicable to the transient phenomena we have observed. The classical expression for the velocity of a vortex ring[14] in terms of its dimensions and vorticity assumes an isolated vortex ring whose cross-sectional diameter is very small compared to its core diameter. Even if a

[13] See particularly reference 8, Plate XIII, photographs 1-4; reference 9, Figs. 4–7, p. 501.

[14] See reference 2, Eq. 163 (7).

correction were made for the finite cross section of the ring, it would appear unreasonable to apply this expression to the initial acceleration phase of the vortex ring we have observed, because during this phase it is certainly not isolated but is abstracting energy from the persistent air jet which formed it. The intrinsic velocity which the ring, at each successive instant, would have if isolated is perhaps superimposed on the translational velocity acquired from the jet. The persistence of the jet appears to prohibit also the application of the simple expression[15] for the strength of a vortex ring produced by a single impulsive pressure, although the origin of the vortex ring may perhaps be understood qualitatively in terms of this concept. Measurement of the vorticity of the ring would be very helpful in determining the extent to which the simple theory obtains during this transient period.

ACKNOWLEDGMENTS

The authors take this opportunity to thank F. T. McClure, H. S. Morton, and R. A. Alpher for their encouragement and support. The theoretical predictions of shock-tube performance were made by Ione D. V. Faro. We are grateful to F. K. Hill and the entire staff of the hypersonic wind tunnel at the Applied Physics Laboratory for their assistance and the use of their facilities and to C. T. Holliday for assistance with the schlieren system. The electronic timer and spark source were made available to us by E. L. Gayhart, R. B. Edmonson, and H. L. Olsen. We are indebted to F. N. Frenkiel, G. M. Edelman, R. G. Prescott, and R. J. Rubin for helpful discussion of the phenomena.

[15] See reference 3, Eq. E-III-(3.4).

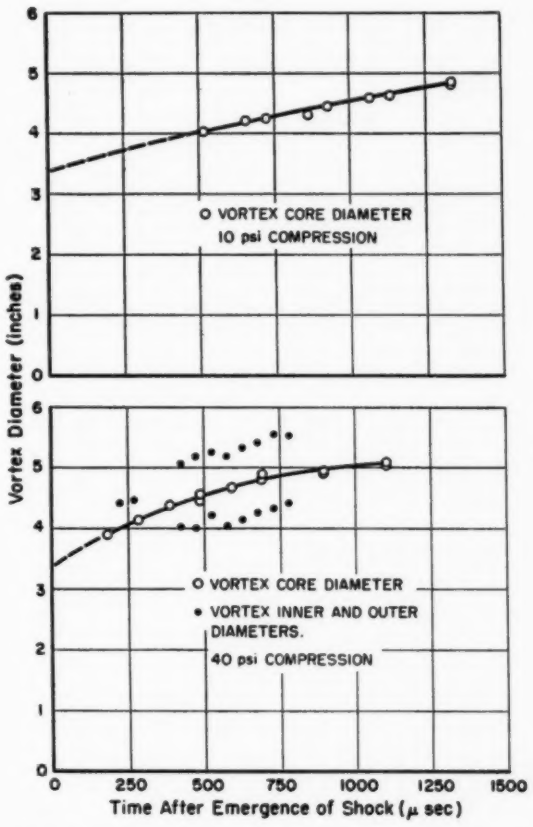


FIG. 6. Diameter of vortex ring as a function of time.

A Property of Bessel Functions and Its Application to the Theory of Two Rheometers*

HERSHEL MARKOVITZ
Mellon Institute, Pittsburgh, Pennsylvania
(Received January 7, 1952)

A method of expanding expressions of the type $J(\xi)Y(\sigma)-J(\sigma)Y(\xi)$ and $J'(\xi)Y(\sigma)-J'(\sigma)Y(\xi)$ is given. Here J and Y are Bessel functions of the first and second kind. These forms occur frequently in physical problems with cylindrical symmetry. The type of expansion used here can be used with other functions that are solutions of second-order differential equations. The use of these expansions allows the development of theories for two rheometers used for determining dynamic mechanical properties of liquids. In these mathematical treatments, the inertia of the sample is taken into account.

I. INTRODUCTION

IN the rheological work of this laboratory two instruments are used to study the response of liquid systems to sinusoidal stresses. The determination of the mechanical properties of the sample from the primary data requires knowledge of the motion of the moving parts of the instrument and of the sample itself. Both instruments have cylindrical symmetry: In one a cylindrical rod is caused to vibrate along its axis in a coaxial tube containing the sample; in the other a cylindrical bob is suspended coaxially in the liquid sample contained in a cup which is oscillated with a definite frequency and amplitude. It is frequent practice in the use of such instruments to so design them, or to assume, that the inertia of the sample has no effect on the motion or can be adequately accounted for in a comparatively simple way. Considerable simplification in mathematical treatment is usually thus achieved. The original designer of the vibrating rod instrument considered this problem and concluded, though an exact treatment was not carried out, that the mass of the sample need not be considered for those conditions for which the instrument was designed. The designers of the prototype for the suspended bob instrument ignored the problem in most of their applications and suggested a rough approximation for cases where the mass of the sample might become important.

Early experience with a modified form of the latter apparatus showed that in a large fraction of the situations to which the instrument was applicable the mass of the sample could not be ignored and that spurious results were obtained from application of the simple equations of motion. An exact treatment of the problem for each instrument was therefore undertaken.

Such an exact analysis is not difficult to perform. However, the solution is of inconvenient form since the equations of motion in each case are Bessel equations and, in the result, the desired mechanical properties are part of the complex argument of Bessel functions. In both cases combinations of the type $J(\sigma)Y(\xi)-J(\xi)Y(\sigma)$ and $J'(\sigma)Y(\xi)-J(\xi)Y'(\sigma)$ appear in the solution. Here J and Y represent Bessel functions of the first and second kinds, respectively, whose arguments σ and ξ are in general complex and contain the unknown quantities sought. The problem thus becomes that of devising a method of freeing the desired quantities from the complex arguments of Bessel functions in the above combinations. This is done by the use of Taylor series expansions, the Wronskian, and other properties of these functions. With this mathematical treatment a restriction on the planning of this kind of apparatus is removed in that the design need no longer be such that the inertia of the sample exert a negligible effect.

The properties of the Bessel functions developed here should be useful in other physical problems. The approach of using Taylor expansions and the Wronskian, as is done here, is of sufficient generality so that it should be applicable to other types of functions which are solutions of second-order differential equations.

II. MATHEMATICAL SECTION

The Bessel functions of the mth order of the first and second kind, usually indicated by $J_m(z)$ and $Y_m(z)$, respectively, are two independent solutions of the differential equation

$$\frac{d^2y}{dz^2}=\left(\frac{m^2}{z^2}-1\right)y-\frac{1}{z}\frac{dy}{dz}. \tag{1}$$

Here z may be complex.

Upon differentiation of both sides, (1) becomes

$$y^{(3)}=\left(-\frac{2m^2}{z^3}\right)y+\left(\frac{m^2+1}{z^2}-1\right)y^{(1)}-\frac{1}{z}y^{(2)}, \tag{2}$$

where, in general, the symbol $y^{(n)}$ stands for the nth derivative of y with respect to z. If (1) is substituted into (2) in order to eliminate $y^{(2)}$, the latter can be written as

$$z^3y^{(3)}=[-3m^2+z^2]y+[(m^2+2)z-z^3]y^{(1)}. \tag{3}$$

On continuing this process one can also find that

$$z^4y^{(4)}=[(m^4+11m^2)-(2m^2+3)z^2+z^4]y+[-(6m^2+6)z+2z^3]y^{(1)}, \tag{4}$$

* This work was supported by the Reconstruction Finance Corporation, Synthetic Rubber Division, at Mellon Institute.

$$z^5y^{(5)}=[-(10m^4+50m^2)+(12m^2+12)z^2-2z^4]y$$
$$+[(m^4+35m^2+24)z-(2m^2+7)z^3+z^5]y^{(1)}, \quad (5)$$

$$z^6y^{(6)}=[(m^6+85m^4+274m^2)-(3m^4+78m^2+60)z^2$$
$$+(3m^2+9)z^4-z^6]y+[-(15m^4+225m^2+120)z$$
$$+(18m^2+33)z^3-3z^5]y^{(1)}. \quad (6)$$

In general, one can write

$$z^ny^{(n)}=(\sum_{j=0}^{\infty}\alpha_{n,2j}z^{2j})y+(\sum_{j=0}^{\infty}\alpha_{n,2j+1}z^{2j+1})y^{(1)}, \quad (7)$$

where $\alpha_{n,s}$ is the coefficient of z^s when $z^ny^{(n)}$ is put in the form (7). From looking at the above samples of this form, it is clear that

$$\alpha_{n,s}=0 \quad s>n \quad \text{and} \quad s<0, \quad (8)$$

$$\alpha_{0,0}=1, \quad (9)$$

$$\alpha_{1,0}=0, \quad \alpha_{1,1}=1. \quad (10)$$

In addition to the cases treated in this paper, many physical problems, especially those with cylindrical symmetry, where conditions are given on two boundaries, lead to solutions where Bessel functions appear in combinations such as $J(\xi)Y(\sigma)-J(\sigma)Y(\xi)$ and $J'(\xi)Y(\sigma)-J(\sigma)Y'(\xi)$. If the functions with argument σ are expanded in a Taylor's series about ξ, the first combination can be written

$$J(\xi)Y(\sigma)-J(\sigma)Y(\xi)$$
$$=\sum_{n=0}^{\infty}(\sigma-\xi)^n[J(\xi)Y^{(n)}(\xi)-J^{(n)}(\xi)Y(\xi)]/n!.$$

Since (7) holds for both J and Y, which are solutions of (1), one can deduce that

$$J(\xi)Y(\sigma)-J(\sigma)Y(\xi)=\sum_{n=0}^{\infty}\frac{1}{n!}(\sigma-\xi)^n\sum_{j=0}^{\infty}\xi^{2j+1-n}$$
$$\times[J(\xi)Y'(\xi)-J'(\xi)Y(\xi)]\alpha_{n,2j+1}$$
$$=\frac{2}{\pi}\sum_{j=0}^{\infty}\xi^{2j}\sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{\sigma-\xi}{\xi}\right)^n\alpha_{n,2j+1}, \quad (11)$$

where the order of summation has been interchanged and the fact[1] that

$$J(\xi)Y'(\xi)-J'(\xi)Y(\xi)=2/\pi\xi \quad (12)$$

has been used. Similar expansions can be made for the second combination mentioned above and for others.

If the value of $(\sigma-\xi)/\xi$ is sufficiently small, the series will converge rapidly and a good estimate of its value can be made by using values of $\alpha_{n,s}$ obtained by inspection of (3) through (6) and perhaps a few higher derivatives. In some cases it is possible to perform the sums exactly. Examples of this for Bessel functions of the zeroth and first orders are shown below.

This type of expansion is especially convenient when the arguments of the Bessel functions are of intermediate size so that one can use neither the power series valid for small arguments nor the asymptotic expansion which is valid for large arguments.

Before proceeding to the evaluation of the sums indicated in (11) it is convenient to obtain certain relations between the coefficients corresponding to two successive derivatives.

If (7) is differentiated the result, after use of (1), is

$$z^{n+1}y^{(n+1)}=[m^2\alpha_{n,1}-n\alpha_{n,0}$$
$$+\sum_{j=0}^{\infty}\{m^2\alpha_{n,2j+3}-(n-2j-2)\alpha_{n,2j+2}-\alpha_{n,2j+1}\}z^{2j+2}]y$$
$$+[\sum_{j=0}^{\infty}\{\alpha_{n,2j}-(n-2j)\alpha_{n,2j+1}\}z^{2j+1}]y^{(1)}.$$

From the definitions of $\alpha_{n+1,k}$, one can then write

$$\alpha_{n+1,2j}=m^2\alpha_{n,2j+1}-(n-2j)\alpha_{n,2j}-\alpha_{n,2j-1}, \quad (13)$$

$$\alpha_{n+1,2j+1}=-(n-2j)\alpha_{n,2j+1}+\alpha_{n,2j}. \quad (14)$$

The above equations are valid for Bessel functions of any order m. We shall now consider the special cases of orders zero and one.

A. Bessel Functions of Order Zero

To avoid confusion, the symbol $a_{j,k}$ will be used to refer to Bessel functions of the zeroth order to correspond to $\alpha_{j,k}$ which refers to the mth order; i.e., in this subsection, we shall put $m=0$ when using the above equations.

For the value $j=0$, (13) becomes

$$a_{n+1,0}=-na_{n,0}$$

with the use of (8). From (9) and (10), it then follows that

$$a_{n,0}=0. \quad n>0. \quad (15)$$

With the substitution $j=0$, (14) becomes

$$a_{n+1,1}=-na_{n,1}$$

with the aid of (15). The solution of this difference equation is

$$a_{n,1}=(-1)^{n-1}(n-1)! \quad n\geq 1. \quad (16)$$

When $j=1$, (13) becomes

$$a_{n+1,2}=-(n-2)a_{n,2}-a_{n,1},$$

[1] G. N. Watson, *A Treatise on the Theory of Bessel Functions* (Cambridge University Press, Cambridge, 1945), p. 76.

from which one can see that

$$a_{3,2} = -a_{2,1} \tag{17_3}$$

$$a_{4,2} = -a_{3,2} - a_{3,1} \tag{17_4}$$

$$a_{5,2} = -2a_{4,2} - a_{4,1} \tag{17_5}$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$a_{j,2} = -(j-3)a_{j-1,2} - a_{j-1,1} \tag{17_j}$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$a_{n,2} = -(n-3)a_{n-1,2} - a_{n-1,1} \quad n \geq 3. \tag{17_n}$$

If (17_3) is multiplied by $(-1)^{n-3}(n-3)!/0!$, (17_4) by $(-1)^{n-4}(n-3)!/1!$, $\cdots(17_j)$ by $(-1)^{n-j}(n-3)!/(j-3)!$, etc., and the resulting $n-2$ equations added together, the result is

$$\sum_{j=3}^{n} (-1)^{n-j}\frac{(n-3)!}{(j-3)!}a_{j,2} = -\sum_{j=3}^{n} (-1)^{n-j}\frac{(n-3)!}{(j-3)!}(j-3)a_{j-1,2} - \sum_{j=3}^{n} (-1)^{n-j}\frac{(n-3)!}{(j-3)!}a_{j-1,1} \quad n \geq 3.$$

This can be reduced to

$$a_{n,2} = -\sum_{j=3}^{n} (-1)^{n-j}\frac{(n-3)!}{(j-3)!}a_{j-1,1} \quad n \geq 3. \tag{18}$$

With the use of (16), (18) becomes

$$a_{n,2} = (-1)^{n+1}(n-3)! \sum_{j=3}^{n} (j-2) = (-1)^{n+1}(n-1)!/2 \quad n \geq 3. \tag{19}$$

By inspection of (1), it is seen that

$$a_{2,2} = -1. \tag{20}$$

Using methods similar to those used in deriving (19), the results

$$\begin{aligned} a_{n,3} &= (-1)^n[(n-1)! + 2(n-3)!]/4 && n \geq 3, \\ a_{n,4} &= (-1)^n[(n-1)! + 4(n-3)!]/16 && n \geq 5, \\ a_{n,5} &= -(-1)^n[(n-1)! + 8(n-3)! + 24(n-5)!]/64 && n \geq 5 \end{aligned} \tag{21}$$

can be obtained.

With the aid of the above formulas, certain useful sums may be evaluated.

From (9) and (15), it is clear that

$$\sum_{n=0}^{\infty} a_{n,0}x^n/n! = 1.$$

By using (8), (16), and the Taylor expansion

$$\ln(1+x) = -\sum_{n=1}^{\infty} (-1)^n x^n/n, \quad |x| < 1, \tag{22}$$

one finds that

$$\sum_{n=0}^{\infty} a_{n,1}x^n/n! = -\sum_{n=1}^{\infty} (-1)^n(n-1)!x^n/n! = -\sum_{n=1}^{\infty} (-1)^n x^n/n = \ln(1+x). \tag{23}$$

Equations (19), (20), and (22) can be used to yield

$$\sum_{n=0}^{\infty} a_{n,2}x^n/n! = \tfrac{1}{2}[\ln(1+x) - x - x^2]. \tag{24}$$

To evaluate the corresponding sum for $a_{n,3}$, it is convenient to notice that integrating both sides of (22) between 0 and x gives

$$(1+x)\ln(1+x) - x = \sum_{n=2}^{\infty} (-1)^n x^n/n(n-1), \tag{25}$$

while a repeated integration reveals that

$$\tfrac{1}{4}[2(1+x)^2\ln(1+x) - 2x - 3x^2] = -\sum_{n=3}^{\infty} (-1)^n x^n/n(n-1)(n-2). \tag{26}$$

With this result and (21) it can be shown that

$$\sum_{n=0}^{\infty} a_{n,3}x^n/n! = [-(2+2x+x^2)\ln(1+x) + 2x + x^2]/4. \tag{27}$$

Also it can be shown that

$$\sum_{n=0}^{\infty} a_{n,5}x^n/n! = [6(6+12x+10x^2+4x^3+x^4)\ln(1+x) - 36x - 54x^2 - 36x^3 - 9x^4]/384. \tag{28}$$

It is clear that the sums in (23)–(28), all of which are based on (22), are valid only for the condition $|x| < 1$. Using the sums here obtained in (11), one obtains the desired expansion.

B. Bessel Functions of Order One

With the substitution $m = 1$ in Eqs. (13) and (14) we can write

$$b_{n+1,0} = b_{n,1} - nb_{n,0} \tag{29_0}$$

$$b_{n+1,1} = -nb_{n,1} + b_{n,0} \tag{29_1}$$

$$b_{n+1,2} = b_{n,3} - (n-2)b_{n,2} - b_{n,1} \tag{29_2}$$

$$b_{n+1,3} = -(n-2)b_{n,3} + b_{n,2} \tag{29_3}$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$b_{n+1,2j} = b_{n,2j+1} - (n-2j)b_{n,2j} - b_{n,2j-1} \tag{29_{2j}}$$

$$b_{n+1,2j+1} = -(n-2j)b_{n,2j+1} + b_{n,2j}, \tag{29_{2j+1}}$$

where the $b_{j,k}$ refer to Bessel functions of the first order and correspond to the $\alpha_{j,k}$ which refer to the mth order. If Eqs. (29) are added together in pairs, the result is

$$b_{n+1,0}+b_{n+1,1} = -(n-1)(b_{n,1}+b_{n,0}) \quad (30_0)$$

$$b_{n+1,2}+b_{n+1,3} = -(n-3)(b_{n,2}+b_{n,3})-b_{n,1} \quad (30_1)$$

$$\vdots \qquad \vdots$$

$$b_{n+1,2j}+b_{n+1,2j+1} = -(n-2j-1)(b_{n,2j}+b_{n,2j+1}) - b_{n,2j-1}. \quad (30_j)$$

Using the symbols $r_n=b_{n,2j}+b_{n,2j+1}$ and $s_n=b_{n,2j-1}$, the prototype equation (30_j) becomes

$$r_{n+1} = -(n-2j-1)r_n - s_n$$

and

$$r_{2j+2} = -s_{2j+1} \quad (31_2)$$

$$r_{2j+3} = -r_{2j+2} - s_{2j+2} \quad (31_3)$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$r_{2j+k} = -(k-2)r_{2j+k-1} - s_{2j+k-1} \quad k\geq 2 \quad (31_k)$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$r_{2j+m} = -(m-2)r_{2j+m-1} - s_{2j+m-1} \quad m\geq 2. \quad (31_m)$$

If (31_2) is multiplied by $(-1)^{m-2}(m-2)!/0!$, (31_3) by $(-1)^{m-3}(m-2)!/1!$, $\cdots$, (31_k) by $(-1)^{m-k}(m-2)!/(k-2)!$, etc., and all the resulting equations added together, the sum is

$$r_{2j+m}=\sum_{k=1}^{m-1}(-1)^{m-k}s_{2j+k}(m-2)!/(k-1)! \quad m\geq 2$$

or

$$b_{2j+m,2j}+b_{2j+m,2j+1} = \sum_{k=1}^{m-1}(-1)^{m-k}\frac{(m-2)!}{(k-1)!}b_{2j+k,2j-1} \quad m\geq 2. \quad (32)$$

If Eqs. (29) are subtracted in pairs and a similar treatment followed, it can be shown that

$$b_{2j+m,2j}-b_{2j+m,2j+1} = (-1)^{m+j}m! - \sum_{k=1}^{m}(-1)^{m-k}m!\,b_{2j+k-1,2j-1}/k! \quad m\geq 1. \quad (33)$$

We can now proceed to the evaluation of the $b_{j,k}$. For $j=0$, (32) and (33) become

$$b_{m,0}+b_{m,1}=0 \qquad m\geq 2,$$

$$b_{m,0}-b_{m,1}=(-1)^m m! \quad m\geq 1,$$

with the use of (8). The solution of these two simultaneous equations is

$$b_{m,0}=(-1)^m m!/2 \quad m\geq 2$$

$$b_{m,1}=-(-1)^m m!/2 \quad m\geq 2. \quad (34)$$

Equations (9) and (10) can be used for $m<2$.

For $j=1$, (32) becomes

$$b_{m+2,2}+b_{m+2,3}=\sum_{k=1}^{m-1}(-1)^{m-k}\frac{(m-2)!}{(k-1)!}b_{k+2,1} = (-1)^{m+1}\frac{(m-2)!}{2}\sum_{k=1}^{m-1}\frac{(k+2)!}{(k-1)!} = (-1)^{m+1}(m+2)!/8 \quad m\geq 2 \quad (35)$$

after the use of (34) and the results of Appendix I to evaluate the sum. If the substitution $m+2=n$ is made, (35) becomes

$$b_{n,2}+b_{n,3}=(-1)^{n-1}n!/8 \quad n\geq 4. \quad (36)$$

For $j=1$, (33) with the use of (34) and Appendix I becomes

$$b_{n,2}-b_{n,3}=(-1)^{n-1}[2(n-2)!+n!]/4 \quad n\geq 3. \quad (37)$$

The result of adding (36) and (37) is

$$b_{n,2}=(-1)^{n-1}[3n!+4(n-2)!]/16 \quad n\geq 4,$$

while subtracting (37) from (36) gives

$$b_{n,3}=(-1)^n[n!+4(n-2)!]/16 \quad n\geq 4.$$

The following useful results can be obtained in a similar fashion:

$$b_{n,5}=(-1)^{n+1}[n!+12(n-2)!+72(n-4)!]/384 \quad n\geq 6,$$

$$b_{n,7}=(-1)^n[n!+24(n-2)!+432(n-4)! +2880(n-6)!]/18432 \quad n\geq 8.$$

For use in (54) it is necessary to obtain the following sums in a way similar to that used in Sec. IIA. These sums are all taken from n equals zero to infinity.

$$\sum b_{n,1}x^n/n! = x(x+2)/2(x+1) \quad (38)$$

$$\sum b_{n+1,1}x^n/n! = [1+(1+x)^{-2}]/2 \quad (39)$$

$$\sum b_{n,3}x^n/n! = [4(1+x)\ln(1+x)-4x-2x^2 -2x^3+x^4(1+x)^{-1}]/16 \quad (40)$$

$$\sum b_{n+1,3}x^n/n! = [4\ln(1+x)+1-6x-3x^2 -(1+x)^{-2}]/16 \quad (41)$$

$$\sum b_{n,5}x^n/n! = [-12(2+4x+3x^2+x^3) \times\ln(1+x)+24x+36x^2+20x^3 +4x^4+2x^5-x^6(1+x)^{-1}]/384 \quad (42)$$

$$\sum b_{n+1,5}x^n/n! = [-12(4+6x+3x^2) \times\ln(1+x)-1+50x+45x^2 +20x^3+5x^4+(1+x)^{-2}]/384 \quad (43)$$

$$\sum b_{n,7}x^n/n! = [24(1+x)(5+10x+9x^2+4x^3+x^4) \times \ln(1+x) - 1 - 119x - 301x^2 - 315x^3 - 175x^4 - 49x^5 - 7x^6 - x^7 + (1+x)^{-1}]/18432 \quad (44)$$

$$\sum b_{n+1,7}x^n/n! = [24(15+38x+31x^2+20x^3+5x^4) \times \ln(1+x) + 1 - 362x - 729x^2 - 604x^3 - 221x^4 - 42x^5 - 7x^6 - (1+x)^{-2}]/18432. \quad (45)$$

III. APPLICATION TO INSTRUMENTS

A. The Coaxially Vibrating Instrument

One of the instruments used is of the type first described by W. Philippoff[2] but bears a greater resemblance to that of T.L. Smith, J. D. Ferry, and F. W. Schremp.[3] The latter consider the question of whether the inertia of the sample can be neglected and use a rough approximation to arrive at an affirmative answer. Here we shall attempt to get a quantitative answer to this problem. Aside from this inclusion, the treatment here follows theirs in many details.

In this apparatus a cylindrical rod of radius r is vibrated along the axis of a coaxial cylindrical vessel (of radius R) which contains the liquid being investigated. A cylindrical liquid lamina of height h, arbitrary radius z, and thickness dz will experience a force due to the shearing at z, equal to $-2\pi rh\eta(\partial V/\partial z)$, where V is the instantaneous velocity in the vertical direction at z, η is the complex dynamic viscosity (sometimes called the mechanical impedance function) and is equal to $\eta'-i(G'/\omega)$, η' is the usual dynamic viscosity, G' is the dynamic rigidity, and ω is 2π times the frequency. At $z+dz$ the shearing force in the lamina will be $2\pi zh\eta(\partial V/\partial z)+[\partial(2\pi zh\eta\partial V/\partial z)/\partial z]dz$. The excess pressure P at the level of the bottom of the rod will exert a force of $2\pi zPdz$ on this element of volume. As a result of these forces the lamina of material will have an acceleration of dV/dt such that

$$2\pi hz\rho(dV/dt)dz = 2\pi zPdz + [\partial(2\pi zh\eta\partial V/\partial z)/\partial z]dz, \quad (46)$$

where ρ is the density of the investigated material.

Since the system is being driven at a frequency of $\omega/2\pi$, we can write $V=v\exp(i\omega t)$ and $P=p\exp(i\omega t)$, where v and p are independent of time. With these definitions of v and p, (46) leads to the differential equation

$$\frac{d^2v}{dz^2}+\frac{1}{z}\frac{dv}{dz}-\frac{i\omega\rho}{\eta}v=-\frac{p}{h\eta}$$

whose solution can be written as

$$v=AJ_0(az)+BY_0(az)+p/(i\omega\rho h), \quad (47)$$

where J_0 and Y_0 are Bessel functions of the zeroth order of the first and second kind, respectively, $a^2=-i\omega\rho/\eta$, and A and B are constants whose values can be determined from the boundary conditions. Since the outer wall at $z=R$ is fixed, $V(R)=v(R)=0$. At $z=r$, the liquid will have the same velocity (V_r) as the vibrating rod. Upon application of the boundary conditions, A and B can be evaluated and (47) becomes

$$-p\{Y_0(ar)J_0(az)-J_0(ar)Y_0(az)-J_0(aR)Y_0(ar) + J_0(ar)Y_0(aR)\}-(i\omega\rho hv_r-p) \times\{Y_0(aR)J_0(az)-J_0(aR)Y_0(az)\} = v(i\omega\rho h)\{J_0(aR)Y_0(ar)-J_0(ar)Y_0(aR)\}. \quad (48)$$

Since the volume of liquid forced down by the rod must be equal to that moving upward in the annular space,

$$\pi v_r r^2+\int_r^R v(2\pi z)dz=0. \quad (49)$$

Before substituting (48) into (49) in order to solve for P, (48) can be expanded by using (11), (23), (27), and (28) of the mathematical section of this paper to obtain the expressions

$$Y_0(ar)J_0(az)-J_0(ar)Y_0(az) = -\frac{2}{\pi}\left[\ln\left(\frac{z}{r}\right)-\frac{i\omega\rho}{4\eta}\left\{-(z^2+r^2)\ln\left(\frac{z}{r}\right)+z^2-r^2\right\} -\frac{\omega^2\rho^2}{128\eta^2}\left\{(2r^4+8z^2r^2+2z^4)\ln\left(\frac{z}{r}\right)-3z^4+3r^4\right\} +O(\rho^3)\right]. \quad (50)$$

It then turns out that

$$p=-4h\eta v_r(1+i\omega\rho\gamma/\eta)/[(R^2+r^2)\ln(R/r)-R^2+r^2],$$

where

$$\gamma=-\frac{R^2+5r^2}{16}+\frac{r^4\ln(R/r)}{4(R^2-r^2)} +\frac{(R^2-r^2)^2\ln(R/r)}{24[(R^2+r^2)\ln(R/r)-R^2+r^2]}$$

and where, as in the later work in this section, terms of higher order than $(i\omega\rho/\eta)$ have been dropped in the expansion.

The experimental quantity which is conveniently determined with this apparatus is the mechanical impedance Z_M which is defined as the ratio of the complex driving force (F) to the velocity (V) of the system. In this apparatus the rod is driven electrically by passing an alternating current through a rigidly attached coil located between the poles of a loud speaker magnet. The rod is held in position by four wires so that for a

[2] W. Philippoff, Physik. Z. 35, 884, 900 (1934).
[3] Smith, Ferry, and Schremp, J. Appl. Phys. 20, 144 (1949).

small displacement s in the vertical direction, it experiences a force of $-ks$. It is also acted on by the forces $2\pi rh\eta(\partial V/\partial z)_r$ due to the shearing action of the liquid and πr^2P from the excess pressure. As a result of all of these forces the rod acquires the acceleration dV_r/dt in the vertical direction such that

$$F-ks+2\pi rh\eta(\partial V/\partial z)_r+\pi r^2P=mdV_r/dt,$$

where m is the mass of the moving system and $V=ds/dt$. Since the force is sinusoidal, we can also write $F=f\exp(i\omega t)$. We thus find that

$$Z_M=F/V_r=f/v_r=i\omega m+k/i\omega -2\pi rh\eta(dv/dz)_r/v_r-\pi r^2p/v_r.$$

By using the results obtained previously in this section, it can be shown that

$$Z_M=i\omega m+k/i\omega+2\pi h\eta(R^2+r^2) \times(1+i\omega\rho\beta/\eta)/[(R^2+r^2)\ln(R/r)-R^2+r^2], \quad (51)$$

where

$$\beta=-r^4/2(r^2+R^2)-r^2\ln(R/r)/2 +(R^2-r^2)^3/24(R^2+r^2)[(R^2+r^2)\ln(R/r)-R^2+r^2].$$

This result is just that previously obtained[3] except for the factor $(1+i\omega\rho\beta/\eta)$. For the apparatus as described by Ferry, $r=0.16$ cm and $R=0.28$ cm, and $\beta=0.00061$ cm^2. Thus as long as the frequencies are not too high and η is not too small, this correction will not be significant. However, if the instrument is to be used for less viscous liquids at higher frequencies or if more precise results are wanted, the full equation (51) must be used.

B. The Torsionally Oscillating Rheometer

Another instrument used in our laboratory has similarities to one described by Goldberg and Sandvik.[4] In this rheometer, a cylindrical bob of radius r is suspended coaxially to a depth h in a liquid contained in a cup of radius R. The cup is caused to oscillate through a very small fixed angular amplitude θ_R. It is desired to find a relationship between the mechanical properties of the sample and the amplitude and phase angle of the motion of the cup relative to that of the suspended cylinder. A paper[5] which has been published recently in another journal contains details of the construction of this apparatus and results obtained with it.

Consider a cylindrical lamina of the liquid of height h with inner radius z and outer radius $z+dz$. The torque acting on the cylindrical surface of the sample at z is $-2\pi z^2\eta hz(\partial\dot\vartheta/\partial z)$, where $\dot\vartheta$ is the angular velocity of the liquid at z. At $z+dz$, the torque on the surface is

$$2\pi z^3\eta h(\partial\dot\vartheta/\partial z)+\{\partial[2\pi z^3h\eta(\partial\dot\vartheta/\partial z)]/\partial z\}dz,$$

and the net torque on the lamina is

$$\{\partial[2\pi z^3h\eta(\partial\dot\vartheta/\partial z)]/\partial z\}dz.$$

As a result, the element of volume will acquire an angular acceleration $\ddot\vartheta$ such that

$$2\pi z^3h\rho\ddot\vartheta dz=\{\partial[2\pi z^3h\eta(\partial\dot\vartheta/\partial z)]/\partial z\}dz. \quad (52)$$

Since the motion is a forced oscillation with frequency $\omega/2\pi$, the substitution $\vartheta=\theta\exp(i\omega t)$ can be made, and (52) can be reduced to

$$d^2\theta/dz^2+(3/z)(d\theta/dz)-i\rho\omega\theta/\eta=0,$$

the solution of which is

$$\theta=[AJ_1(az)+BY_1(az)]/z, \quad (53)$$

where J_1 and Y_1 are Bessel functions of the first order, A and B are arbitrary constants to be determined by the boundary conditions, and a is $(-i\omega\rho/\eta)^{\frac{1}{2}}$.

At the surface of the cup containing the liquid, the motion can be represented by $\vartheta(R)=\theta_R\exp(i\omega t)$ since the liquid is following the forced motion of the cup which has the amplitude θ_R. The suspended cylinder experiences a shearing torque equal to $2\pi r^2h\eta(\partial\dot\vartheta/\partial r)_r$ due to the liquid, and a torque $-k\vartheta_r$ due to the suspension wire which has a torsion constant k. As a result the bob acquires an angular acceleration such that

$$2\pi h\eta r^2(\partial\dot\vartheta/\partial r)_r-k\vartheta_r=I\ddot\vartheta_r,$$

where I is the moment of inertia of the cylinder. Putting the boundary conditions into (53) one can arrive at the equation of motion for the inner cylinder

$$\theta_r=\frac{ab\theta_Rr^2[J_1'(ar)Y_1(ar)-J_1(ar)Y_1'(ar)]}{(1+br^2)[Y_1(ar)J_1(aR)-J_1(ar)Y_1(aR)]-abr^3[J_1(aR)Y_1'(ar)-J_1'(ar)Y_1(aR)]}, \quad (54)$$

where $b=-i2\pi h\eta\omega/(I\omega^2-k)$ and $J_1'(ar)$ is the value of $dJ_1(az)/d(az)$ at $z=r$. Although (54) represents the exact relationship between the desired quantity η and the measured quantities θ_r/θ_R, it is in an inconvenient form since η is buried in the complex argument of the Bessel functions.

However, making a Taylor expansion of the functions with argument ar about aR as indicated in (11), and

[4] H. Goldberg and O. Sandvik, Anal. Chem. 19, 123 (1947).

[5] Markovitz, Yavorsky, Harper, Zapas, and De Witt, Rev. Sci. Instr. 23, 430 (1952).

using (12) and (38)–(45), it then follows that

$$\frac{\theta_R}{\theta_r}=1+\frac{i}{\mathfrak{y}}\left[\frac{I\omega^2-k}{4\pi h\omega}\frac{R^2-r^2}{R^2r^2}+\frac{\omega\rho}{8}\frac{(R^2-r^2)^2}{R^2}\right]$$
$$-\frac{1}{\mathfrak{y}^2}\left[\frac{(I\omega^2-k)\rho}{32\pi h}\left(4\ln\frac{r}{R}+\frac{R^2}{r^2}-\frac{r^2}{R^2}\right)\right.$$
$$\left.+\frac{\omega^2\rho^2}{192R^2}\left\{(R^2-r^2)(R^4-5R^2r^2-2r^4)+12R^2r^4\ln\frac{R}{r}\right\}\right]$$
$$+\frac{i}{\mathfrak{y}^3}\left[\frac{(I\omega^2-k)\omega\rho^2}{768\pi hR^2r^2}\left\{12R^2r^2(R^2+r^2)\ln\frac{R}{r}\right.\right.$$
$$\left.-(R^2-r^2)(R^4+10R^2r^2+r^4)\right\}$$
$$-\frac{\omega^3\rho^3}{9216R^2}\left\{24R^2r^4(3R^2+2r^2)\ln\frac{R}{r}\right.$$
$$\left.\left.+(R^2-r^2)(R^6-11R^4r^2-47R^2r^4-3r^6)\right\}\right]+\cdots.$$

It turns out that in all cases studied in this laboratory, the term in $1/\mathfrak{y}^3$ and all higher terms can be neglected, but that terms lower than this make an appreciable contribution. As a result $\mathfrak{y}$ can be obtained from the experimental quantities by solving a quadratic equation with complex coefficients. In this connection the tables[6] for $(1+ix)^{\frac{1}{2}}$ for $0<x<1$ are convenient to use.

The author wishes to thank Dr. T. W. DeWitt and Dr. John R. Bowman for checking the manuscript and for helpful discussions.

APPENDIX I

We seek to evaluate sums of the type

$$\sum_{k=n}^{m}\frac{k!}{(k-n)!}.$$

For this purpose it is convenient to examine the equality

$$S\equiv\sum_{k=0}^{m}x^k=\frac{1-x^{m+1}}{1-x},$$

which when differentiated n times yields

$$S^{(n)}\equiv\frac{d^nS}{dx^n}\equiv\sum_{k=n}^{m}\frac{k!}{(k-n)!}x^{k-n}$$
$$=\sum_{j=0}^{n}\left[\frac{n!}{j!(n-j)!}\left\{\frac{d^j}{dx^j}(1-x^{m+1})\right\}\right.$$
$$\left.\times\left\{\frac{d^{m-j}}{dx^{m-j}}\left(\frac{1}{1-x}\right)\right\}\right]=\frac{n!}{(1-x)^{n+1}}\left[1-x^{m+1}\right.$$
$$\left.-\sum_{j=1}^{n}\left\{\frac{(m+1)!}{j!(m+1-j)!}(1-x)^jx^{m+1-j}\right\}\right] \quad (A1)$$

since[7]

$$\frac{d^n yz}{dx^n}=\sum_{j=0}^{n}\frac{n!}{j!(n-j)!}\frac{d^jy}{dx^j}\frac{d^{n-j}z}{dx^{n-j}},$$

$$\frac{d^j}{dx^j}(1-x^{m+1})=-\frac{(m+1)!}{(m+1-j)!}x^{m+1-j},$$

$$\frac{d^j}{dx^j}\left(\frac{1}{1-x}\right)=\frac{j!}{(1-x)^{j+1}}.$$

If the substitution $x=1-t$ is made in (A1) and the limit taken as t approaches zero, the left-hand side immediately becomes

$$\sum_{k=n}^{m}\frac{k!}{(k-n)!}, \quad (A2)$$

which is the sum we seek. Applying the same operation to the right side yields

$$S^{(n)}=\frac{n!}{t^{n+1}}\left[1-\sum_{j=0}^{n+1}(-1)^j\frac{(m+1)!t^j}{(m+1-j)!j!}\right.$$
$$\left.-\sum_{j=1}^{n}\frac{(m+1)!t^j}{j!}\sum_{i=0}^{m+1-j}(-1)^i\frac{t^i}{(m+1-j-i)!i!}\right] \quad (A3)$$

on use of the binomial expansion. After the new dummy index

$$k=i+j$$

is substituted in the last term of (A3), the rearranged result can be written

$$S^{(n)}=\frac{n!(m+1)!}{t^{n+1}}\left[-\sum_{k=1}^{n}\frac{(-1)^kt^k}{(m+1-k)!k!}\right.$$
$$\times\left(1+\sum_{j=1}^{k}\frac{(-1)^jk!}{j!(k-j)!}\right)-\sum_{k=n+1}^{m+1}\frac{(-1)^kt^k}{(m+1-k)!k!}$$
$$\left.\times\left(1+\sum_{j=1}^{n}\frac{(-1)^{k-j}}{j!(k-j)!}\right)\right],$$

which, due to the fact that

$$[1+(-1)]^k=0=\sum_{j=0}^{k}\frac{k!}{j!(k-j)!}=1+\sum_{j=1}^{k}\frac{(-1)^jk!}{j!(k-j)!},$$

[6] E. Jahnke and F. Emde, *Tables of Functions with Formulas and Curves* (Dover Publications, Inc., New York, 1943), p. 15 of the Addenda.

[7] See for example: B. O. Peirce, *A Short Table of Integrals* (Ginn and Company, Boston, 1929), Eq. 858.

becomes

$$S^{(n)}=\frac{n!(m+1)!}{t^{n+1}}\sum_{k=n+1}^{m+1}\frac{(-1)^{k+1}t^k}{(m+1-k)!k!}\times\left(1+\sum_{j=1}^{n}\frac{(-1)^{k-j}}{j!(k-j)!}\right). \quad \text{(A4)}$$

When the limit of (A4) is taken as t approaches zero, one finds that

$$\lim_{t\to 0}S^{(n)}=(-1)^n\frac{(m+1)!}{(m-n)!(n+1)!}\left(1+\sum_{j=1}^{n}\frac{(-1)^{n+1-j}}{j!(n+1-j)!}\right),$$

which, when the observation is made that

$$[1+(-1)]^{n+1}=0=\sum_{j=0}^{n+1}\frac{(n+1)!(-1)^j}{(n+1-j)!j!}$$

$$=(n+1)!\left[1+(-1)^{n+1}+\sum_{j=1}^{n}\frac{(-1)^{n+1-j}}{j!(n+1-j)!}\right],$$

becomes simply

$$\frac{(m+1)!}{(n+1)(m-n)!}=\sum_{k=n}^{m}\frac{k!}{(k-n)!}$$

by using (A2). This is the desired sum.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 10 OCTOBER, 1952

The Concept and Measurement of Impedance in Periodically Loaded Wave Guides*

EDWIN T. JAYNES
Microwave Laboratory, Stanford University, California

(Received April 2, 1952)

A generalization of ordinary circuit theory which enables one to define impedances in any periodic structure is developed, based on the concept of expansion of electromagnetic fields in terms of a set of linearly independent basis fields. Techniques for measurement of impedances in a periodic structure are described, involving a determination of the parameters of a coupling system by an extension of the well-known nodal shift method.

I. INTRODUCTION

IN a coupling system connecting two different transmission lines or wave guides† it is usually desired that a matched line on the load side be reflected through the coupling system as a match in the input transmission line. The obvious way to test for this is to install a matched load and then measure the input SWR. However, an accurate match may be very difficult to obtain; or if the output transmission line happens to be of some special construction, such as the periodically disk-loaded wave guides used in particle accelerators,[1] there may be no independent method of determining when the load line is matched, other than looking at it through some coupling system.

It is always easy, however, to produce a pure reactance (i.e., completely reflecting) termination of the load line, and if the coupling system is lossless, its input impedance must also be a pure reactance. This reactance is determined merely by locating the position of the extremely sharp node on the input transmission line, so that if we can find a way of determining the coupling system parameters from the relation of load reactance to input reactance, we can expect this method to give us a more convenient as well as more accurate determination of the coupling system performance. Once the parameters of the coupling system are known, the relations between its input and output impedance may be used backwards to measure impedances in the load transmission system.

In Sec. II we review the well-known theory of the nodal shift method as applied to "smooth" transmission lines. Sections III, IV, and V are devoted to developing the concept of impedance in periodically loaded wave guides, and in Sec. VI the theory of the nodal shift method is extended to such guides. The resulting measurement technique was developed and used by the writer in 1946, in connection with the design of coupling systems for the Stanford linear accelerator tubes.[1]

II. NODAL SHIFT THEORY

First, we note the familiar fact that the voltage-current relationships of any linear, passive, four-

FIG. 1. Voltage and current conventions for a four-terminal network.

* This work was supported in part by Navy Contract N6-onr-25116.

† We will understand the term "transmission line" to mean any structure in which waves may be propagated according to the usual transmission-line equations. The term thus includes wave guides, coaxial lines, and two-wire lines.

[1] W. W. Hansen, *Consiglio Nazionale de Ricerche* (Tipografia del Senato, Roma, 1948), p. 111; Ginzton, Hansen, and Kennedy, Rev. Sci. Instr. **19**, 89 (1948); E. L. Chu and W. W. Hansen, J. Appl. Phys. **18**, 996 (1947); **20**, 280 (1949); E. L. Chu, Ph.D. thesis (Stanford University, 1951).

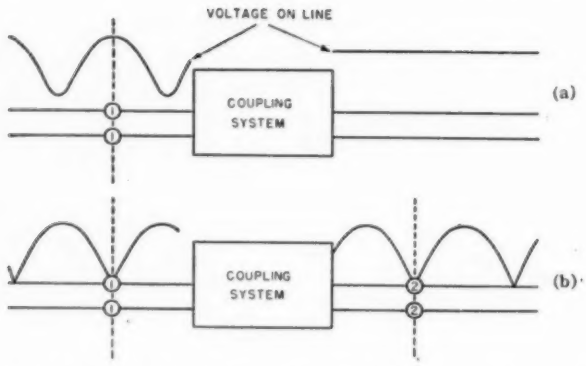


FIG. 2. Definition of reference planes. (a) The load line is matched; (b) the load is a pure reactance.

terminal network may be expressed in terms of the network parameters A, B, C, D, in the form (see Fig. 1 for notation)

$$\begin{aligned} V_1 &= AV_2 + BI_2, \\ I_1 &= CV_2 + DI_2, \\ \begin{vmatrix} A & B \\ C & D \end{vmatrix} &= 1. \end{aligned} \tag{1}$$

For a lossless network, A and D are real, while B and C are imaginary.[2] The input impedance $Z_1 = V_1/I_1$ is thus given in terms of the load impedance $Z_2 = V_2/I_2$ by

$$Z_1 = (AZ_2 + B)/(CZ_2 + D). \tag{2}$$

We may consider the network of which the coupling system is a part as including arbitrary lengths of input and output transmission lines. We now show that for a lossless system the reference planes (terminals) may be chosen so as to simplify (2). Assume that voltages and currents are so defined that all impedances are normalized, i.e., $Z = 1$ on any matched transmission line; and consider the two situations depicted in Fig. 2. In Fig. 2(a) the output line is matched. Choose terminals (1) at a voltage maximum on the input line. Then the impedance Z_1 is real and greater than unity. This particular value of impedance is numerically equal to the voltage SWR η seen on line 1 when line 2 is matched, and may be called the SWR of the coupling system. From Eq. (2) we have

$$\eta = (A + B)/(C + D).$$

In Fig. 2(b) we have chosen the output terminals so that a short there also places a short at the input terminals. For this case, Eq. (2) reduces to

$$B/D = 0, \quad \text{or} \quad B = 0.$$

But since η is real, we must have $C = 0$ also. Therefore $\eta = A/D$, and the general impedance relation (2) reduces, for this choice of terminals, to

$$Z_1 = \eta Z_2. \tag{3}$$

In general, the input impedance to a shorted transmission line of length x is $jZ_0 \tan\theta$, with $\theta = 2\pi x/\lambda$ the electrical length of the line. With normalized impedances, $Z_0 = 1$. From this we see that if a short is placed a distance θ_2 to the right of terminals (2), there will be a short a distance θ_1 to the right of terminals (1) (Fig. 3), where

$$\tan\theta_1 = \eta \tan\theta_2. \tag{4}$$

If one places a movable short in the load transmission line and measures the position of the short in the input line as a function of the position of the load short, Eq. (4) will enable one to find from these data the value of η and the positions of the two reference planes. Usually these are the only parameters of the coupling system in which one is interested.

In practice one would plot the difference $(\theta_1 - \theta_2)$ as a function of θ_2, or if more convenient, the distances $(x_1 - x_2)$ and x_2 may be plotted directly. Since the reference points are not in general known at this stage of the process, one measures position of the nodes from arbitrary origins, and the origin of the $(\theta_1 - \theta_2)$ *vs* θ_2 graph remains arbitrary. The shape and orientation of the plot, however, are determined by the value of η. The equation represented is

$$\begin{aligned} \theta_1 - \theta_2 &= \tan^{-1}\eta \tan\theta_2 - \tan^{-1}\tan\theta_2 \\ &= \tan^{-1}\frac{(\eta - 1)\tan\theta_2}{1 + \eta\tan^2\theta_2}. \end{aligned} \tag{5}$$

As a function of θ_2 this reaches a maximum of

$$\Delta\theta = (\theta_1 - \theta_2)_{\max} = \tan^{-1}(\eta^{\frac{1}{2}} - 1/\eta^{\frac{1}{2}}) \tag{6}$$

at a value of θ_2 given by

$$\tan^2\theta_2 = 1/\eta. \tag{7}$$

Solving (6) for η, we have

$$\eta = (1 + \sin\Delta\theta)/(1 - \sin\Delta\theta), \tag{8}$$

and the reflection coefficient is simply

$$\Gamma = (\eta - 1)/(\eta + 1) = \sin\Delta\theta. \tag{9}$$

Figure 4 is a plot of actual data taken at a wavelength of about 10 cm, with a coupling system of reflection

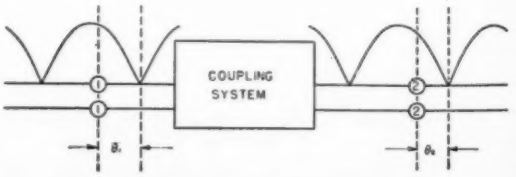


FIG. 3. Definitions of node positions θ_1 and θ_2.

[2] E. A. Guillemin, *Communication Networks* (John Wiley and Sons, Inc., New York, 1935), Vol. II.

coefficient $\Gamma=0.466$. One finds $\Delta\theta$ from the difference between maximum and minimum values of (x_1-x_2) and from this determines Γ and η. Then, as is seen from differentiating (5) at $\theta_2=0$, the positions of the input and output reference points are determined as the point where (x_1-x_2) is midway between the maximum and minimum values, with positive slope. Thus, the three parameters of the coupling system are all determined from a single plot of data.

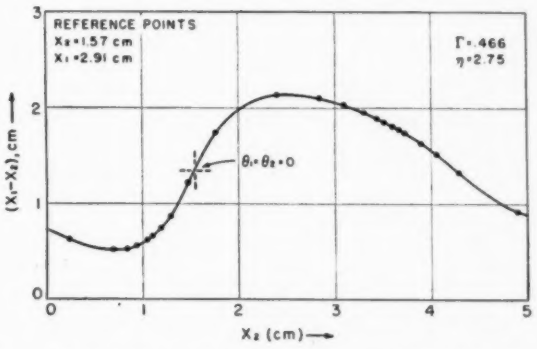


FIG. 4. Experimental plot of Eq. (5). As explained in Sec. VI, these data were actually taken with a periodically loaded guide as the output line.

III. THE CONCEPT OF GENERALIZED IMPEDANCE

When we come to the case of a periodically loaded wave guide (see Fig. 6 for an example), it is not immediately obvious what is meant by impedance and position of a node. In fact, the exact distinction between a running wave and a standing wave is none too clear. One thing which we need is a generalization of the concept of impedance which goes beyond the familiar generalization to wave impedance in "smooth" wave guides (i.e., wave guides in which all boundary conditions are independent of the z coordinate). The concept of impedance is useful because it gives us a certain piece of information about the configuration of electromagnetic fields, and any quantity, however defined, which conveys the same information may be called a generalized impedance.

In any guide in which only one mode type is excited, there are only two linearly independent possible fields; any type of wave, standing or running, may be expressed as a linear combination of any two linearly independent fields. This situation exists in the guides under consideration in regions far from the local fields due to coupling systems, and it is in these regions that we must be able to define what we mean by impedance or node position. Let $(\mathbf{E}_1, \mathbf{H}_1)$ and $(\mathbf{E}_2, \mathbf{H}_2)$ be two different possible field configurations; then any field is specified by two complex amplitudes a_1, a_2:

$$\begin{aligned}\mathbf{E}&=a_1\mathbf{E}_1+a_2\mathbf{E}_2,\\ \mathbf{H}&=a_1\mathbf{H}_1+a_2\mathbf{H}_2.\end{aligned}\tag{10}$$

Instead of giving a_1, a_2, we might specify the product (a_1a_2) and the ratio (a_1/a_2); the former quantity is essentially a measure of the power level, while the latter specifies the field configuration to within a multiplicative constant and is therefore a generalized impedance.

The quantities a_1, a_2 are evidently very loosely analogous to voltage and current. This analogy may be greatly strengthened by a proper choice of the basis fields $(\mathbf{E}_1, \mathbf{H}_1)$ and $(\mathbf{E}_2, \mathbf{H}_2)$. In the first place, we should like to have the power flow given by the usual formula

$$P=\tfrac{1}{2}Re(V\bar{I})=\tfrac{1}{2}Re(a_1\bar{a}_2),\tag{11}$$

where the bar denotes the complex conjugate. This power flow is given by an integral of the complex Poynting vector across any surface S' spanning the guide; using the expansion (10) we have

$$\begin{aligned}P&=\tfrac{1}{2}Re\int_{S'}(\mathbf{E}\times\bar{\mathbf{H}})\cdot d\mathbf{S}\\ &=\tfrac{1}{2}Re\int_{S'}[|a_1|^2(\mathbf{E}_1\times\bar{\mathbf{H}}_1)+a_1\bar{a}_2(\mathbf{E}_1\times\bar{\mathbf{H}}_2)\\ &\qquad+\bar{a}_1a_2(\mathbf{E}_2\times\bar{\mathbf{H}}_1)+|a_2|^2(\mathbf{E}_2\times\bar{\mathbf{H}}_2)]\cdot d\mathbf{S}.\end{aligned}\tag{12}$$

This will of course be independent of which particular surface S' is chosen, provided losses in the guide can be neglected. Now if the equality of (11) and (12) is to be an identity in a_1, a_2, it is evident that we must have

$$Re\int_{S'}(\mathbf{E}_1\times\bar{\mathbf{H}}_1)\cdot d\mathbf{S}=Re\int_{S'}(\mathbf{E}_2\times\bar{\mathbf{H}}_2)\cdot d\mathbf{S}=0,\tag{13}$$

whereupon (12) becomes

$$P=\tfrac{1}{2}Re\left\{a_1\bar{a}_2\int_{S'}(\mathbf{E}_1\times\bar{\mathbf{H}}_2+\bar{\mathbf{E}}_2\times\mathbf{H}_1)\cdot d\mathbf{S}\right\}.\tag{14}$$

In (14) use has been made of the fact that any term in (12) may be replaced by its complex conjugate since only the real part is taken. Equation (13) implies that we should choose basis fields that carry no power; i.e., standing waves for which $\mathbf{E}$ and $\mathbf{H}$ are everywhere $\pi/2$ out of phase. The basis fields $(\mathbf{E}_1, \mathbf{H}_1)$, $(\mathbf{E}_2, \mathbf{H}_2)$ thus correspond to two different pure reactance terminations or to two different node positions. They still contain arbitrary phase factors, however, and in order to reduce (14) to (11) we should choose them so that the integral in (14), which is in any event a constant independent of the surface S', becomes real. Proper choices of the amplitudes of the basis fields will then give us

$$\int_{S'}[\mathbf{E}_1\times\bar{\mathbf{H}}_2+\bar{\mathbf{E}}_2\times\mathbf{H}_1]\cdot d\mathbf{S}=1.\tag{15}$$

If we choose the following conditions (sufficient but

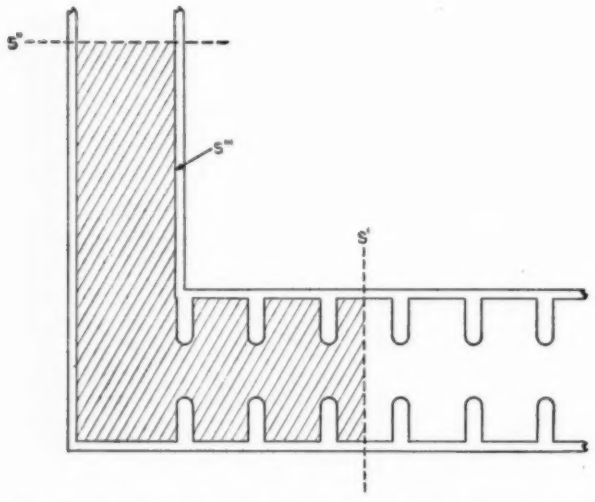


FIG. 5. Region of integration in Eq. (20). The smooth wave guide at the top is coupled to the periodically loaded one through a wide iris.

not necessary):

$$\begin{aligned}&\mathbf{E}_1, \mathbf{H}_2 \text{ real},\\ &\mathbf{E}_2, \mathbf{H}_1 \text{ imaginary},\end{aligned} \tag{16}$$

then (13) is automatically satisfied, and with proper normalization so is (15). Thus if the conditions (15), (16) are imposed on the basis fields, the power flow is given correctly by (11).

It remains to be demonstrated that a_1, a_2 are still analogous to voltage and current in the sense that the voltage and current (V, I) seen at some other part of a complicated network (in particular in a smooth transmission line on the other side of some coupling system to which the periodically loaded guide is connected) are related to a_1 and a_2 by the equations appropriate to a four-terminal network. Merely from the fact that Maxwell's equations are linear, we know that a relation of the form

$$\begin{pmatrix} V \\ I \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \tag{17}$$

must exist; it must now be shown that the determinant of the transformation is unity:

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = 1, \tag{18}$$

and that in the case of a lossless network A and D are real while B and C are pure imaginary. The first of these follows from the reciprocity theorem; let us consider the functions $(\mathbf{E}_1, \mathbf{H}_1)$, $(\mathbf{E}_2, \mathbf{H}_2)$ as standing not only for the base functions in the periodically loaded guide, but also their analytic continuations through the coupling system and into the smooth transmission line on which V, I are measured. Then at all points inside the complete network both of these fields will satisfy Maxwell's equations

$$\nabla\times\mathbf{H} = j\omega\epsilon\mathbf{E} \quad \nabla\times\mathbf{E} = -j\omega\mu\mathbf{H}, \tag{19}$$

in which in general ϵ, μ are functions of position. If the coupling network is lossy, ϵ will be complex at some points within it. Because of (19) it is easily shown that

$$\nabla\cdot(\mathbf{E}_1\times\mathbf{H}_2 - \mathbf{E}_2\times\mathbf{H}_1) = 0$$

or

$$\oint_S (\mathbf{E}_1\times\mathbf{H}_2 - \mathbf{E}_2\times\mathbf{H}_1)\cdot d\mathbf{S} = 0, \tag{20}$$

where we integrate over any closed surface. Equation (20) is an expression of the reciprocity theorem in the form given by Lorentz. Let us now carry out this integration over the surface S consisting of the surface S' spanning the periodically loaded guide, as used in Eqs. (12)–(15), the metallic walls S''' of the system extending from S' through the coupling system, and out to a plane surface S'' normal to the smooth guide on the other side of the coupling system, as shown in Fig. 5. S'' is the reference plane at which the voltage and current V, I are measured. The region enclosed by this closed surface $S = S' + S'' + S'''$ is shaded in Fig. 5.

Because of the boundary conditions satisfied by $\mathbf{E}_1, \mathbf{E}_2$ on the metallic surface, the integral over S''' vanishes separately. Then, if the positive normals to S', S'' are chosen so that both point in the direction of power flow, we have

$$\int_{S''} (\mathbf{E}_1\times\mathbf{H}_2 - \mathbf{E}_2\times\mathbf{H}_1)\cdot d\mathbf{S} = \int_{S'} (\mathbf{E}_1\times\mathbf{H}_2 - \mathbf{E}_2\times\mathbf{H}_1)\cdot d\mathbf{S}.$$

But with the choice of phases given by (16) the integral over S' is the same as that in (15), which we have normalized to unity. (This incidentally gives an independent proof that the integral (15) is independent of our choice of the surface S'.) Therefore,

$$\int_{S''} (\mathbf{E}_1\times\mathbf{H}_2 - \mathbf{E}_2\times\mathbf{H}_1)\cdot d\mathbf{S} = 1. \tag{21}$$

Now at the reference plane S'' the voltage and current are ordinarily defined as proportional to the transverse electric and magnetic fields, respectively. The transverse fields at S'' are therefore expressible as

$$\begin{aligned}\mathbf{E}(S'') &= V\mathbf{E}_t(S''),\\ \mathbf{H}(S'') &= I\mathbf{H}_t(S''),\end{aligned}$$

where $\mathbf{E}_t$, $\mathbf{H}_t$ are normal mode functions so normalized that power flow is given by (11). We consider the smooth transmission system to support only a single propagating mode at the frequency of operation and the reference plane S'' to be far from the local fields of the coupling system. Therefore,

$$\int_{S''} (\mathbf{E}\times\mathbf{H})\cdot d\mathbf{S} = VI. \tag{22}$$

If the smooth transmission system is a coaxial line, it is easily shown that (22) is valid where V and I are the actual voltage between conductors and actual current flow, respectively. Thus we can put in general

$$\int_{S''} (\mathbf{E}_1\times\mathbf{H}_2-\mathbf{E}_2\times\mathbf{H}_1)\cdot\mathbf{dS}=V_1I_2-V_2I_1=1, \quad (23)$$

where V_1 is the voltage at S'' when the field $\mathbf{E}_1$ is excited, I_2 is the current when $\mathbf{H}_2$ is excited, etc. But reference to (17) shows that

$$V_1=A, \quad I_1=C,$$
$$V_2=B, \quad I_2=D,$$

since, for example, A and C are the voltage and current when $a_1=1$, $a_2=0$, i.e., when the basis field $(\mathbf{E}_1, \mathbf{H}_1)$ is excited. Therefore (23) is equivalent to (18). Furthermore, it is evident that when the entire network is lossless, the choice of phases made in (16) for the region of the periodically loaded guide applies also to the analytical continuation of $\mathbf{E}_1, \mathbf{H}_1$, etc., to the surface S'', so that A and D are real; B and C are imaginary. Thus all of the properties of the network equations (1) which were needed for deriving the nodal shift formulas have been shown to hold for the connection (17) in which a_1, a_2 are used in place of voltage and current.

IV. CHANGE OF BASIS FIELDS

The conditions which we have imposed on the basis fields are merely that E_1, E_2 be linearly independent and that (15), (16) be satisfied. Even with these very weak restrictions, the analogy between the expansion coefficients (a_1, a_2) and the circuit quantities of voltage and current has become so complete that all of the relations of general circuit theory remain valid when (a_1, a_2) are used. The basis fields are, however, far from being uniquely determined by (15), (16); if (E_1, H_1) and (E_2, H_2) is any set allowed by our conditions, then an equally good set is (E_1', H_1'), (E_2', H_2'), where

$$\begin{aligned} E_1'&=G_{22}E_1-G_{21}E_2, & H_1'&=G_{22}H_1-G_{21}H_2, \\ E_2'&=-G_{12}E_1+G_{11}E_2, & H_2'&=-G_{12}H_1+G_{11}H_2, \end{aligned} \quad (24)$$

in which the matrix

$$G=\begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix}$$

has determinant unity, G_{11} and G_{22} being real while G_{12} and G_{21} are pure imaginary. These are just the conditions that the new basis fields should satisfy (15), (16) when the old ones do. A matrix with these properties will be called a G matrix. A field which has expansion coefficients (a_1, a_2) with respect to the old basis fields will have coefficients (a_1', a_2') with respect to the new ones, where

$$\begin{pmatrix} a_1' \\ a_2' \end{pmatrix}=\begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \quad (25)$$

for we then find that

$$\mathbf{E}=a_1\mathbf{E}_1+a_2\mathbf{E}_2=a_1'\mathbf{E}_1'+a_2'\mathbf{E}_2', \text{ etc.}$$

The product of any two G matrices is another G matrix, so that the totality of all transformations of the type (25) forms an infinite group. Since the matrix

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

of general circuit parameters for a lossless four-terminal network is a G matrix, we see that there is a one-to-one correspondence between all possible lossless four-terminal networks and possible changes of basis fields which preserve the analogies of (a_1, a_2) to voltage and current. Any such change of basis fields is equivalent to looking at the original voltage and current through some particular lossless network.

Since a G matrix has three independent parameters, we may expect in general that three more independent conditions may be imposed on our basis fields in addition to the above ones before they become uniquely determined. Additional conditions which would be convenient in various problems are that the two basis fields should represent the same stored energy, that the generalized impedance (a_1/a_2) should equal unity when a pure running wave is excited, etc.

V. APPLICATION TO A PERIODICALLY LOADED GUIDE

Suppose that the wave guide contains regularly spaced structures with a repetition distance "a"; in other words, the entire system is invariant under translations in the z direction through the distance a. The guide may be considered as a cascaded line of "unit cells" each of length a, but the boundaries of the cells are arbitrary. Now we can choose any two linearly independent basis fields $\mathbf{E}_1(x, y, z)$; $\mathbf{E}_2(x, y, z)$ which satisfy conditions (15) and (16). Suppressing the x, y coordinates for brevity, we write the expansion of a general field in the periodic structure as

$$\mathbf{E}(z)=a_1\mathbf{E}_1(z)+a_2\mathbf{E}_2(z). \quad (26)$$

We now inquire under what conditions $\mathbf{E}(z)$ represents a pure running wave. By this we mean that the value of $\mathbf{E}$ in one cell is a multiple of its value in an adjacent one:

$$\mathbf{E}(z+a)=e^{-j\beta}\mathbf{E}(z), \quad (27)$$

where $j\beta$ is the propagation constant per cell, which is purely imaginary in the passband. This corresponds to the condition of impedance match in the theory of iterated networks, if we consider each cell to correspond

to a four-terminal network.‡ To see this, we note that the expansion (26), while valid at all values of z, may be considered as applying to one section only, say the n'th, in which case a_1, a_2 correspond to voltage and current at the output of the n'th of the iterated networks. The fields in the $(n+1)$th section may be expanded in terms of new basis fields $\mathbf{E}_1'$, $\mathbf{E}_2'$ which bear the same relation to the $(n+1)$th section as $\mathbf{E}_1$, $\mathbf{E}_2$ did to the n'th; in other words, they are obtained by translating $\mathbf{E}_1$, $\mathbf{E}_2$ one section down the pipe:

$$\begin{aligned}\mathbf{E}_1'(z)&=\mathbf{E}_1(z-a),\\ \mathbf{E}_2'(z)&=\mathbf{E}_2(z-a),\end{aligned} \tag{28}$$

and the new expansion is

$$\mathbf{E}(z)=a_1'\mathbf{E}_1'(z)+a_2'\mathbf{E}_2'(z) \tag{29}$$

so that a_1', a_2' correspond to voltage and current delivered by the $(n+1)$th network. This is a particular type of change of basis fields as discussed above, since $\mathbf{E}_1'$, $\mathbf{E}_2'$ are linearly related to $\mathbf{E}_1$, $\mathbf{E}_2$ through some G matrix:

$$\begin{pmatrix}\mathbf{E}_1\\ \mathbf{E}_2\end{pmatrix}=\begin{pmatrix}G_{11} & G_{12}\\ G_{21} & G_{22}\end{pmatrix}\begin{pmatrix}\mathbf{E}_1'\\ \mathbf{E}_2'\end{pmatrix}. \tag{30}$$

We now substitute these relations into the condition (27) for a running wave. Express $\mathbf{E}(z+a)$ by means of Eqs. (28) and (29) and $\mathbf{E}(z)$ by Eq. (26), and it assumes the form

$$\mathbf{E}(z+a)=a_1'\mathbf{E}_1(z)+a_2'\mathbf{E}_2(z)=e^{-j\beta}[a_1\mathbf{E}_1(z)+a_2\mathbf{E}_2(z)]$$

or, since $\mathbf{E}_1$, $\mathbf{E}_2$ are linearly independent,

$$a_1'=e^{-j\beta}a_1, \quad a_2'=e^{-j\beta}a_2 \tag{31}$$

$$a_1'/a_2'=a_1/a_2. \tag{32}$$

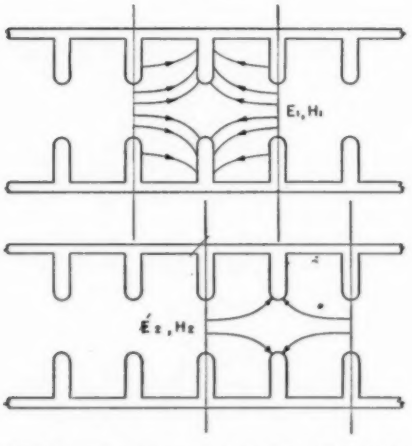


FIG. 6. Electric lines of a possible choice of basis fields for the disk-loaded linear accelerator pipe.

‡ We are, however, particularly anxious to avoid drawing equivalent circuits of the sections, because one of the principal purposes of this analysis is to show that a rigorous treatment of fields leads to relations of the same form as the ordinary circuit-theory relations independently of any approximate equivalent circuit.

The impedance is the same in each cell for a running wave, the same situation that one finds in the theory of iterated networks. As in the classical filter theory, we expect that the propagation constant $j\beta$ and the characteristic impedance (32) are determined by the frequency and the structure of the network. The only properties of the network that are needed to determine them are the elements of the G matrix in (30). This is clear from the relation (25) which, using (31), becomes

$$e^{-j\beta}\begin{pmatrix}a_1\\ a_2\end{pmatrix}=\begin{pmatrix}G_{11} & G_{12}\\ G_{21} & G_{22}\end{pmatrix}\begin{pmatrix}a_1\\ a_2\end{pmatrix}. \tag{33}$$

Thus, $e^{-j\beta}$ is an eigenvalue, while a_1, a_2 are the components of the corresponding eigenvector, of the G matrix connecting the original basis fields with their translations through one period. By standard methods, we obtain the explicit solutions

$$e^{-j\beta}=\tfrac{1}{2}\{(G_{11}+G_{22})\pm[(G_{11}+G_{22})^2-4]^{\frac{1}{2}}\}, \tag{34}$$

$$a_1/a_2=(e^{-j\beta}-G_{22})/G_{21}. \tag{35}$$

The two values of β from (34) satisfy the relation $\beta'=-\beta''$, corresponding to the two directions of propagation. In the passbands β is real, and we have

$$\cos\beta=\tfrac{1}{2}(G_{11}+G_{22})<1. \tag{36}$$

We now observe that relations (34), (35) may be greatly simplified by a particular choice of basis fields. We may choose $\mathbf{E}_1$ and $\mathbf{E}_2$ to differ essentially only by a translation of one period down the guide, provided that the fields so defined are linearly independent. (In general, this will be the case, but the edges of the passbands are critical frequencies at which $\mathbf{E}_1(z)=\pm\mathbf{E}_1(z+a)$ so that linear independence fails.) In addition, we must introduce a 90° phase shift as required by condition (16). Therefore, let us define E_2 by

$$\mathbf{E}_2(z)=-j\mathbf{E}_1(z-a)=-j\mathbf{E}_1'(z). \tag{37}$$

A possible choice of these fields for the Stanford linear accelerator is sketched in Fig. 6. Then the G matrix in (30) becomes

$$G=\begin{pmatrix}G_{11} & -j\\ -j & 0\end{pmatrix}, \tag{38}$$

while (34), (35), and (36) reduce to

$$e^{-j\beta}=G_{11}/2\pm[(G_{11}/2)^2-1]^{\frac{1}{2}}, \tag{39}$$

$$(a_1/a_2)=je^{-j\beta}, \tag{40}$$

$$\cos\beta=G_{11}/2. \tag{41}$$

The characteristic impedance and propagation constant are not essentially different quantities with this choice of basis fields. At the edges of the passband, $\beta=0$ or π, and the characteristic impedance becomes purely imaginary; while at the center of the passband, $\beta=\pi/2$, and the characteristic impedance is unity.

To sum up the results found above, there is a complete analogy between wave propagation in a periodic

structure and the classical theory of filters composed of iterated four-terminal networks. The device of expanding fields in terms of a set of linearly independent basis fields has enabled us to find coefficients a_1, a_2 which satisfy the same relations as do voltage and current in lumped-constant network theory. This is not just a coincidence, for ordinary circuit theory is a special case of the analysis presented here. In any lumped-constant four-terminal network we can define fields $(\mathbf{E}_1, \mathbf{H}_1)$ to be those resulting from unit voltage, zero current at the output (i.e., open-circuited load), while $(\mathbf{E}_2, \mathbf{H}_2)$ are the fields resulting from unit current, zero voltage at the same terminals. Then any possible state of excitation of the network is described by the expansion (26) in which a_1, a_2 are the output voltage and current in the ordinary sense. The G matrix giving the rule of translation of basis fields through one section then reduces to the matrix of general circuit parameters A, B, C, D as treated by Guillemin.[2] Thus, by introducing the concept of basis fields and then allowing ourselves a greater freedom of choice in their definitions than is usual (the extent of this freedom being precisely described by the group of G matrices), we create a rather sweeping generalization of ordinary circuit-theory, which applies without approximation to any periodic structure.

VI. EXPERIMENTAL MEASUREMENTS

We are now able to say just what it is that we want a coupling system to the periodically loaded guide to do; when load conditions are so adjusted that we have a pure running wave in it, as in the Stanford linear accelerator, we want a match to appear also in the smooth transmission line on the other side of the coupling system. The problem is to find an experimental procedure, using only pure reactance terminations of the loaded wave guide, that will tell us what the input impedance on the smooth transmission line would be if there were a running wave in the loaded guide.

We may produce all reactive values of the impedance (a_1/a_2) in the loaded guide by sliding a metal shorting-plug into it at various distances, as shown in Fig. 7. However, we do not at the outset know how this reactance is related to the position of the shorting plug. This may be found from the behavior of the impedance seen on the other side of the coupling system as follows. First we define the input reference plane and the basis fields in a manner analogous to the definitions of Sec. II:

(a) Match the periodic guide: $a_1/a_2 = j\exp(-j\beta)$, and choose the input reference plane at a point on the input line where the impedance seen looking toward the coupling system is $V/I = Kj\exp(-j\beta)$, with K real. If β is very small (i.e., the repetition distance a is very small compared to a wavelength), such a point may not exist, in which case we redefine the cell to contain n of the small cells, $n=2, 3, 4, \cdots$. This multiplies β by n, and a suitable value of n can always be found such that this reference plane will exist.

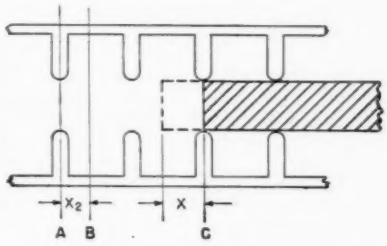


FIG. 7. Metallic plunger for varying load reactance, with definitions of quantities used in Fig. 9. The actual plunger is a brass tube with fingers that ensure good metallic contact when a disk is reached.

(b) Find the plunger position which places a voltage node at this input reference plane, and define the fields in the periodic guide for this plunger position (and sufficiently far from the local fields of the coupling system and plunger) as $(\mathbf{E}_2, \mathbf{H}_2)$.

Then we have, remembering that A and D are real and B and C are imaginary,

$$\text{from (b):}\quad B=0,$$
$$\text{from (a):}\quad C=0,\quad K=A/D;$$

and the impedance relation, for this choice of reference plane and basis fields, again reduces to a simple proportionality

$$V/I = K(a_1/a_2). \tag{42}$$

By transmission-line theory, the SWR of the coupling system is then

$$\eta = \tfrac{1}{2}(K+K^{-1})\csc\beta + [\tfrac{1}{4}(K+K^{-1})^2\csc^2\beta - 1]^{\frac{1}{2}}, \tag{43}$$

which, for $\beta=\pi/2$, reduces to $\eta=K$.

As in Sec. II, the above are *definitions* of reference plane and basis fields, not experimental methods for finding them, because by hypothesis we have no experimental way of knowing when the periodic guide is matched until the parameters of some coupling system have been found. We can find these parameters as follows: First, note that when the plunger is moved in through one repetition period (as defined in (a) above), the basis fields $\mathbf{E}_1, \mathbf{H}_1$ are then excited so that there is an infinite impedance at the input reference plane; experimentally, pushing the plunger in through the distance a moves the node in the smooth transmission line just a quarter-wavelength. Suppose we measure the node position x_1 on the input line as a function of the plunger position x and plot a graph of $x_1(x)-x_1(x+a)$ *versus* x. At the point where this crosses $(\lambda/4)$, x_1 is the position of the input reference plane, x is the plunger position which creates the basis fields $(\mathbf{E}_2, \mathbf{H}_2)$, and $(x+a)$ is the plunger position which creates $(\mathbf{E}_1, \mathbf{H}_1)$. (Note that according to Fig. 7, the direction of increasing x is opposite to the direction of increasing z as used in Sec. V. The reason for this is experimental convenience.)

The analogy between the relations of this section and those of Sec. II may be strengthened if we define an effective node position θ_2 in the periodic guide for any plunger position by

$$j\tan\theta_2 = (a_1/a_2).$$

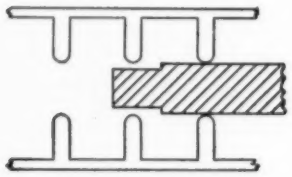


FIG. 8. Modification of plunger to avoid contact discontinuities.

Since, as before, the node position on the input line is given by $j\tan\theta_1=V/I$, Eq. (42) assumes a form analogous to (4):

$$\tan\theta_1=K\tan\theta_2. \tag{44}$$

The remaining quantities which we wish to find are the effective node position θ_2 (or what is more convenient experimentally, the length $x_2=a\theta_2/\beta$) as a function of plunger position x, the impedance magnification constant K, and the phase-shift per section, β. To find these, let us express the fields as functions of the longitudinal coordinate z, and the plunger position x:

$$\mathbf{E}(z,x)=a_1(x)\mathbf{E}_1(z)+a_2(x)\mathbf{E}_2(z).$$

If we pull the plunger out by a distance a_1, the new fields are just translated in the $+z$ direction by an amount a:

$$\mathbf{E}(z,x-a)=a_1(x-a)\mathbf{E}_1(z)+a_2(x-a)\mathbf{E}_2(z)$$
$$=\mathbf{E}(z-a,x)=a_1(x)\mathbf{E}_1(z-a)+a_2(x)\mathbf{E}_2(z-a).$$

From this, using (30), we obtain

$$\begin{pmatrix}a_1(x-a)\\ a_2(x-a)\end{pmatrix}=\begin{pmatrix}0 & j\\ j & G_{11}\end{pmatrix}\begin{pmatrix}a_1(x)\\ a_2(x)\end{pmatrix}. \tag{45}$$

Now calling the plunger position $x=0$ at the position which generates the basis fields $(\mathbf{E}_2, \mathbf{H}_2)$, we find that

$$(a_1/a_2)_{x=0}=0 \text{ (definition)},$$
$$(a_1/a_2)_{x=a}=\infty \text{ (in agreement with the previous discussion)},$$
$$z\,(a_1/a_2)_{x=a}=j/G_{11}, \tag{46}$$
$$(a_1/a_2)_{x=2a}=jG_{11}/(G_{11}^2-1). \tag{47}$$

Taking the product and ratio of (46), (47) we have, using the relation $j\tan\theta_1=V/I$ and (42),

$$\frac{(\tan\theta_1)_{x=a}}{(\tan\theta_1)_{x=2a}}=1-\frac{1}{G_{11}^2}, \tag{48}$$
$$(\tan\theta_1)_{x=a}\cdot(\tan\theta_1)_{x=2a}=K^2/(1-G_{11}^2), \tag{49}$$

from which $G_{11}=2\cos\beta$ and K may be determined. Once K is known, we can use Eq. (42) backwards to find the effective node position $x_2(x)$ from the data on x_1 *versus* x already taken. Thus, once more, from an experimental determination of x_1 as a function of x, we can deduce the SWR of the coupling system [using Eq. (43)], the propagation constant of the periodic guide, the position of the reference planes at which the impedance relations are particularly simple, and the effective node position in the periodic guide as a function of plunger position. Most of the work at Stanford has been done at that frequency for which $\beta=\pi/2$, where the situation is greatly simplified, as is seen from inspection of the above equations for that case.

The use of a metallic plunger with a flat end as in Fig. 7 was found to have a disadvantage that when metallic contact is made to a disk in pushing it forward, there is a small but annoying discontinuity in the function $x_1(x)$. However, one can shape the plunger as in Fig. 8, and it is possible to find a length for the extension (1.3 cm in our case) at which no measurable discontinuity in $x_1(x)$ is found. The condition for this is, of course, that there should be no rf voltage across the gap just before it closes. A theoretical reason why the extension should produce this condition has not been sought; we merely record it as an experimental fact that it is not at all difficult to find the proper length.

The function $x_2(x)$ is a universal one, depending on the periodic guide, the frequency, and the plunger, but not on the particular coupling system with which it was obtained. Therefore, once one has this universal curve (Fig. 9 applies to the present linear accelerator guide in use at Stanford) one can return to the method of Sec. II and obtain the parameters of a number of different coupling systems by plotting (x_1-x_2) *versus* x_2. The data of Fig. 4 were taken in this way.

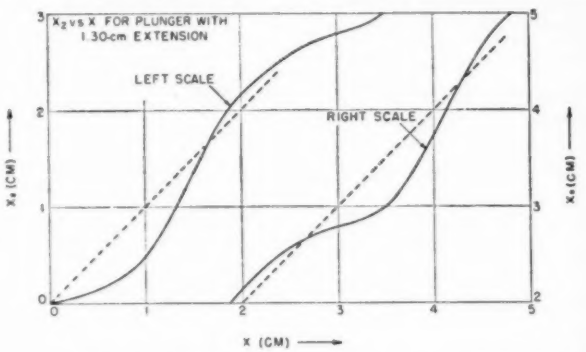


FIG. 9. Effective node position as a function of plunger position for the Stanford linear accelerator guides.

One tricky point which nearly everyone finds disconcerting at first sight concerns the direction of motion of the effective node position. According to Fig. 9 and the definitions in Fig. 7, moving the plunger to the *left* moves the effective node to the *right*, whereas one naïvely expects the opposite behavior. The direction of node motion depends on the details of the fields around the plunger, and one can show that the general criterion is as follows: Push the plunger in through an infinitesimal distance and consider the fields in the small volume that was removed in front of the plunger. If before the movement there was more energy stored in electric than in magnetic fields in this volume, then the effective node moves in the opposite direction to the plunger, and vice versa.

The writer is indebted to the late Dr. W. W. Hansen, under whose direction the experimental part of this work was done, for many helpful discussions.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 10 OCTOBER, 1952

The Complete Solution of the One-Velocity Diffusion Problem at Intermediate and Large Distances

LEWI TONKS
Knolls Atomic Power Laboratory, General Electric Company, Schenectady, New York
(Received March 10, 1952)

An approximate evaluation of the branch-cut integral which occurs in the rigorous solution of the Boltzmann equation for intermediate and large distances is of importance in shielding problems. Such an evaluation leads to Eq. (20) for the point source and Eq. (24) for the plane source. Equation (20) is actually 23 percent high for zero absorption ($f=0$) and $r\sigma=7$, but the error decreases with increasing $r\sigma$ and f. It is seen that the branch-cut term is not simply the contribution of noncollided neutrons.

Figure 2 is a plot of the region in the $f-r\sigma$ plane where the two terms are comparable.

ONE of the simple and important problems of neutron diffusion is the calculation of neutron flux at some distance from an isotropic monoenergetic source in a medium which scatters and absorbs them but does not change their energy. The result of such a calculation has two types of application. It can be used for the diffusion of thermal neutrons which enter the thermal region by moderation from higher energies and whose absorption is a true capture by the nuclei of the diffusing medium. It also can be used for the diffusion of fast neutrons which may come directly from fission and whose "absorption" is their moderation out of the fast energy region by elastic collisions and inelastic scattering.

An important practical application in the latter category is to hydrogenous shields where a single impact with a hydrogen nucleus can be viewed as an "absorption" and a certain fraction of impacts with other heavier nuclei, which may be present, result in "absorption" through inelastic collisions. Such a medium has high "absorption" relative to scattering and we are interested in large attenuations. Both factors tend to emphasize the importance of a term in the solution of the diffusion problem which is not expressible in easily interpretable form and is usually omitted as negligible. This is the definite integral in Eq. (1) below.

In this paper this term is evaluated in algebraic form in a range of the variables which is useful in shielding problems and to a sufficient accuracy for the usual purposes of such problems. Also, a direct comparison is made of the relative importance of this term and the commonly used simple exponential term in the range of the variables where the two are comparable.

The rigorous solution of the Boltzmann equation for the one-velocity case of a point source of neutrons of unit strength in an infinite isotropic medium with isotropic scattering is[1]

$$\phi(r)=\frac{\sigma}{2\pi r}\left\{\frac{\nu^2(1-\nu^2)}{g(\nu^2-f)}e^{-\nu r\sigma}+\left(\frac{1}{2}\right)\int_1^\infty \frac{e^{-\eta r\sigma}d\eta}{\left[1-\frac{g}{\eta}\tanh^{-1}(1/\eta)\right]^2+[\pi g/(2\eta)]^2}\right\}, \tag{1}$$

where $\phi(r)$ is the flux at distance r, f is the ratio of capture to total cross section σ_c/σ, g is the ratio of scattering to total cross section σ_s/σ and

$$(1/\nu)\tanh^{-1}\nu=1/(1-f)=1/g. \tag{2}$$

Approximate solutions of Eq. (2) are

$$\text{for small } f: \quad \nu^2\simeq 3f[1-(4/5)f], \tag{3A}$$

$$\text{for large } f: \quad \nu\simeq 1-2e^{-2/g}-2\frac{4-g}{g}e^{-4/g}. \tag{3B}$$

At $f=0.45$

Eq. (3A) gives $\nu\simeq 0.931$;
Eq. (3B) gives $\nu\simeq 0.939$;
actually $\nu=0.9355$.

Unfortunately, these approximations are poor for evaluating the first bracketed term of the right member of Eq. (1) near $f=0.5$ because of the $(1-\nu^2)$ factor but nearer either zero or unity they are better.

$$\frac{\nu^2(1-\nu^2)}{g(\nu^2-f)}\simeq\begin{cases}\dfrac{3}{2}\dfrac{[1-(4/5)f][1-3f+(12/5)f^2]}{g[1-(6/5)f]}; & \text{small } f \quad (4A)\\[2ex] \dfrac{4e^{-2/g}}{g^2}\left(1+2\dfrac{4-g}{g}e^{-2/g}\right); & \text{large } f. \quad (4B)\end{cases}$$

The ranges for the approximations and the values of the quantities mentioned are shown in Table I.

In the evaluation of the second term in the bracket of the right member of Eq. (1), that is, the branch-cut term, it is convenient in Table II to tabulate the

[1] Marshak, Brooks, and Hurwitz, Nucleonics 4 (5), 10 (1949).

TABLE I. Values of ν and diffusion term coefficient.

f	ν		$1-\nu$	$\frac{\nu^2(1-\nu^2)}{g(\nu^2-f)}$
	Eq. (3A)	Eq. (2)	Eq. (3B)	
0.0	0.00			1.500
0.1	0.525			1.264
0.2	0.710			1.027
0.3	0.827	0.827		0.8045
0.4		0.9075		0.5705
0.45		0.9355		0.4670
0.50		0.9576		0.3650
0.55		0.973		0.2827
0.60		0.9856	0.0143	0.1857
0.70			0.002586	0.0581
0.80			9.09×10^{-5}	0.004549
0.90			4.12×10^{-9}	8.24×10^{-7}

values of

$$k=\eta^{-1}\tanh^{-1}(\eta^{-1})=1/(2\eta)\ln[(\eta+1)/(\eta-1)].$$

Also in Table II is tabulated the coefficient $C(g,\eta)$ of the exponential in the integrand for a series of values of f. As η approaches unity, i.e.,

$$\eta=1+x \tag{5}$$

with $x\ll1$

$$k\simeq\frac{x}{4}+\frac{1-x}{2}\ln(2/x). \tag{6}$$

This family of $C(g,\eta)$ curves, shown in Fig. 1, has the following characteristics:

1. All the curves approach unity as η approaches infinity and lie within 0.05 of unity for $\eta>3$.
2. For $f<1-8\pi^2=0.1895$ the curves lie wholly below unity.
3. As f goes from 0.1895 to $1-4/\pi^2=0.5948$ the curves lie further and further above unity at larger values of η, but not at small values.
4. As f increases further the curves cling closer to unity at all but small values of η, where they reach ever higher maxima at ever smaller η's.
5. These maxima, of values approximately $4/\pi^2g^2$, occur at $\eta\simeq1+2e^{-2/g}$.

We are interested in intermediate and large distances from the source, say

$$\rho\equiv r\sigma>7, \tag{7}$$

so that, judging by the exponential factor under the integral, the pertinent range for x probably does not extend very much beyond 1/7. Thus, unless $C(g,\eta)$ is increasing there more rapidly than approximately e^{7x}, we can confine attention in the branch line integration to, say,

$$0\leq\eta-1=x<e^{-1}=0.368.$$

Using Eq. (5), an obvious expansion using

$$1+x\simeq e^{x}-x^2/2$$

gives for the second bracketed term in Eq. (1)

$$\text{B.T.}\simeq\frac{e^{-\rho}}{2}\int_0^{1/e}\frac{e^{(2-\rho)x}(1-x^2e^{-x})dx}{\left[\frac{g}{2}\left(\ln\frac{2}{x}+\frac{x}{2}\right)-1-x^2\right]+\left[\frac{\pi g^2}{2}\right]}. \tag{8}$$

We now neglect all x's except in $\ln(2/x)$ and in $e^{(2-\rho)x}$;

$$\text{B.T.}\simeq\frac{2e^{-\rho}}{g^2}\int_0^{e^{-1}}\frac{e^{(2-\rho)x}dx}{\left(\ln\frac{2}{x}-\frac{2}{g}\right)^2+\pi^2}. \tag{9}$$

The substitution

$$y=\ln\frac{2}{x}-\frac{2}{g} \tag{10}$$

simplifies the denominator and results in

$$\text{B.T.}\simeq\frac{4}{g^2}e^{-\rho}\int_{1+\ln2-2/g}^{\infty}\times\frac{\exp[-y-2/g-2(\rho-2)e^{-y-2/g}](dy)}{y^2+\pi^2}.$$

Then, to simplify the numerator, the substitution

$$w=y+\frac{2}{g}-\ln(2\rho-4) \tag{11}$$

leads to

$$\text{B.T.}\simeq\frac{2e^{-\rho}}{g^2(\rho-2)}\int_{1-\ln(\rho-2)}^{\infty}\frac{\exp(-w-e^{-w})}{y^2+\pi^2}dw. \tag{12}$$

The numerator of the integrand is a special case of the integrand of the "after-effect function," whose integral from minus to plus ∞ is unity.[2] To find the effect of the denominator we shall first perform the indicated integration by a Gaussian approximation and then use the correction factor arising from the denominator to correct the exact integration of the numerator alone.

TABLE II. Values of k and $C(g,\eta)$.

$\log_{10}\frac{(\eta-1)}{2}$	η	k	$C(g,\eta)$ $f=0$	$f=0.2$	$f=0.5$	$f=0.8$	$f=0.9$
0.978	20	0.0025	0.998	1.000	1.000	1.000	1.0004
0.653	10	0.01003	0.995	1.000	1.004	1.003	1.004
0.301	5	0.0405	0.982	1.0013	1.015	1.012	1.007
0	3	0.1153	0.949	0.9997	1.046	1.036	1.021
−0.301	2	0.274	0.875	0.996	1.112	1.090	1.050
−0.602	1.5	0.537	0.763	0.973	1.234	1.189	1.104
−1.000	1.2	1.00	0.584	0.88	1.475	1.411	1.209
−1.301	1.1	1.382	0.458	0.76	1.653	1.652	1.310
−1.602	1.05	1.77	0.353	0.623	1.746	1.974	1.430
−2.301	1.01	2.62		0.364	1.43	3.092	1.758
−3	1.002	3.44			0.984	5.110	2.198
−4	1.0002	4.60			0.433	9.524	3.162
−5	1.0_42					8.224	4.883
−6	1.0_52						8.313
−7	1.0_62						16.041
−8	1.0_72						32.35
−9	1.0_82						38.49
−10	1.0_92						21.03

[2] Jahncke and Emde, *Tables of Functions III*, p. 5.

Since the lower integration limit was set only by a guess at the range of validity of our approximation, it can be abandoned for minus infinity if this does not make the approximation too crude. This we shall try.

We write the integrand as $e^{-f(w)}$ with

$$f(w)=w+e^{-w}+\ln(y(w)^2+\pi^2), \tag{13}$$

$$f'(w)=1-e^{-w}+\frac{2y}{\pi^2+y^2}, \tag{14}$$

$$f''(w)=e^{-w}+2\frac{\pi^2-y^2}{(\pi^2+y^2)^2}. \tag{15}$$

We denote the root of $f'(w)=0$ by w_0. To a first approximation $w_0=0$, and more exactly

$$w_0\simeq-\ln\left(1+\frac{2y_0}{\pi^2+y_0^2}\right) \tag{16}$$

with (from Eq. (11))

$$y_0\simeq\ln(2\rho-4)-2/g. \tag{17}$$

The approximate integration gives

$$\text{B.T.}\simeq\frac{2(2\pi)^{\frac{1}{2}}e^{-\rho}\exp(-f(w_0))}{g^2(\rho-2)[f''(w_0)]^{\frac{1}{2}}},$$

$$\text{B.T.}\simeq\frac{2(2\pi)^{\frac{1}{2}}e^{-\rho-1}}{g^2(\rho-2)(\pi^2+y_0^2)}\left(1+2\frac{y_0}{\pi^2+y_0^2}\right)^{-\frac{1}{2}}. \tag{18}$$

It is readily seen that a primary effect of the Gaussian approximation is the introduction of the factor $(2\pi)^{\frac{1}{2}}/e$ in place of unity—we therefore replace this factor by unity—and the denominator in (12) contributes the factors containing y_0, which we therefore retain. We have, then,

$$\text{B.T.}\simeq\frac{2e^{-\rho}}{g^2(\rho-2)(\pi^2+y_0^2)}\left(1-\frac{y_0}{\pi^2+y_0^2}\right) \tag{19}$$

with y_0 given by Eq. (17).

Thus Eq. (1) becomes

$$\phi(r)\simeq\frac{\sigma}{2\pi r}\left[\frac{\nu^2(1-\nu^2)}{g(\nu^2-f)}e^{-r\sigma\nu}+\frac{2(1-y_0/(\pi^2+y_0^2))}{g^2(\rho-2)(\pi^2+y_0^2)}e^{-r\sigma}\right] \tag{20}$$

when $\rho>7$.

In the case of an infinite plane source of unit strength, the flux $N(z)$ at a distance z from the source is[1]

$$N(z)=\frac{\nu^2(1-\nu^2)}{g(\nu^2-f)}e^{-z\nu\sigma}+\frac{1}{2}\int_1^\infty\frac{e^{-\eta z}d\eta/\eta}{\left(1-\frac{g}{\eta}\tanh^{-1}\frac{1}{\eta}\right)^2+\left(\frac{\pi g}{2\eta}\right)^2}. \tag{21}$$

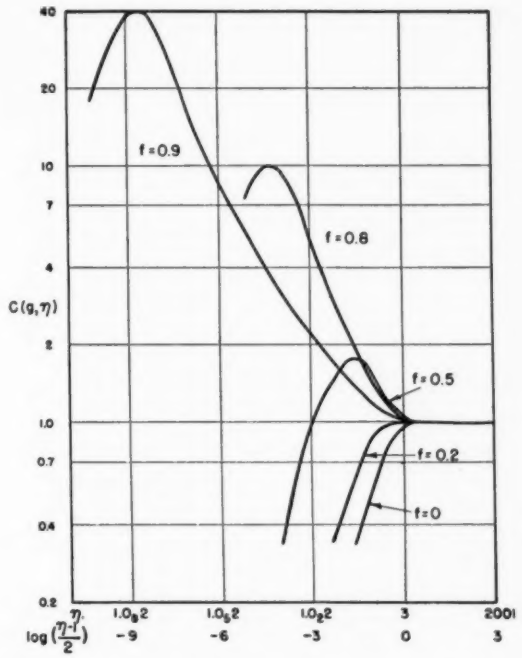


FIG. 1. $C(g,\eta)$ as a function of η, the Laplace inversion variable, for various absorption/scattering ratios.

The same integration methods as before can be used, and it turns out that the only difference is that where $\rho-2$ appeared explicitly before it is to be replaced by $\rho-1$. Thus

$$(\text{B.T.-plane})\simeq\frac{2e^{-\rho}}{g^2(\rho-1)(\pi^2+v_0^2)}\left(1-\frac{v_0}{\pi^2+v_0^2}\right) \tag{22}$$

with

$$v_0\simeq\ln(2\rho-2)-2/g. \tag{23}$$

Thus Eq. (21) becomes

$$N(z)\simeq\frac{\nu^2(1-\nu^2)}{g(\nu^2-f)}e^{-z\sigma\nu}+\frac{2[1-v_0/(\pi^2+v_0^2)]}{g^2(\rho-1)(\pi^2+v_0^2)}e^{-z\sigma} \tag{24}$$

when $\rho>7$.

The consistency of the second term of this with Eq. (19) through the point-plane relation is readily verified if the difference between y_0 and v_0 is neglected and the variations of these quantities with distance are also neglected. To transform from plane-to-point source we have

$$-\frac{d}{\rho d\rho}\left(\frac{e^{-\rho}}{\rho-1}\right)=\frac{e^{-\rho}}{\rho(\rho-1)}\left(1+\frac{1}{\rho-1}\right)\simeq\frac{e^{-\rho}}{\rho(\rho-2)},$$

which is the dependence found for the point source.

To investigate the accuracy of Eq. (19) we note first that from Eqs. (10) and (17) the value x_0 of x corresponding to y_0 is $1/(\rho-2)$. Second, the exponential in the integrand of Eq. (12) shows that the effective range of w, defined by a reduction of the integrand by a factor of about e^{-2} from the maximum, is $-1.2<w<3$,

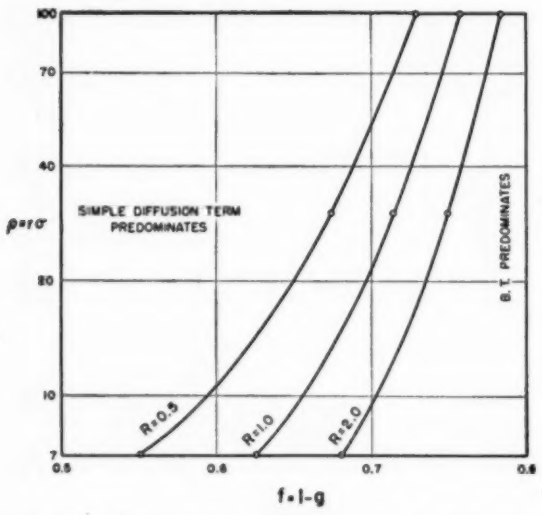


FIG. 2. Region in $r\sigma$—f (i.e., distance—absorption) plane where ratio R of branch-cut term to diffusion term is not far from unity.

which entails a range in x from x_0e^{-3} to $x_0e^{1.2}$. For $\rho=7$, $x_0=0.2$, and the range of x is from 0.01 to 0.66. The approximations leading to Eq. (9) are thus seen to be rather crude. For $\rho=7$ and $g=1$ numerical integration of the branch term as given in Eq. (1) in the form

$$\text{B.T.}=\frac{e^{-\rho}}{2g^2}\int_{w=-\infty}^{\infty}\frac{x(2+2x)^2e^{-\rho x}dw}{\left(\dfrac{2+2x}{g}-\ln(2+x)+\ln x\right)^2+\pi^2} \tag{25}$$

gives

$$2(\text{B.T.})e^{\rho}=0.0316$$

compared to 0.039 from Eq. (19), a 23 percent discrepancy. This will be less for larger values of ρ because $x_0(\rho-2)=1$ and it will be less for smaller values of g because, from Eqs. (16) and (17)

$$w_0\simeq 2\frac{2/g-\ln(2\rho-4)}{\pi^2+[2/g-\ln(2\rho-4)]^2}, \tag{26}$$

so that w_0 increases and x_0 decreases for smaller g.

A naive approach is to regard B.T. as arising from those neutrons which have made no collision. In that case for the point source we should have

$$(\text{B.T.})_{nc}=\frac{e^{-\rho}}{2\rho}, \tag{27}$$

so that

$$\frac{(\text{B.T.})_{nc}}{(\text{B.T.})}\simeq\frac{\rho-2}{\rho}g^2(\pi^2+y_0^2) \tag{28}$$

$$\simeq(1-2/\rho)\{2.47g^2+0.25[g\ln(\rho-2)-1.3]^2\}. \tag{29}$$

Accordingly the true value of B.T. at large ρ is greater or less than the naive according as g is small (large absorption) or near unity. It thus bears little relation to the uncollided neutrons. If the naive view were correct, then $C(g,\eta)$ would be constant at unity.

We have now to compare the relative magnitudes of the diffusion and branch terms. Referring to Eq. (20) we note that $1-y_0/(\pi^2+y_0^2)$ has extreme values so close to unity—namely, 0.84 and 1.16—that we neglect the discrepancy. A little analysis shows that for $\rho>7$ equality of the two terms occurs in the range of f for which Eqs. (3B) and (4B) are valid, and that actually the last approximation can be omitted. We then obtain for the ratio of the two terms in Eq. (20):

$$R=\frac{\text{B.T.}}{\text{Diff.}}=\frac{\exp(2/g-2\rho e^{-2/g})}{2(\rho-2)(\pi^2+y_0^2)}.$$

Figure 2 is a plot of constant R in the $r\sigma-f$ plane for $r\sigma>7$ and $R=0.5$, 1.0, and 2.0. It shows that at a constant $r\sigma$ increasing absorption favors the importance of the branch cut term and that in a particular absorbing medium, i.e., for constant f, the diffusion term predominates at the greater depths, as is well known.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 10 OCTOBER, 1952

The Condensation of Nitrogen in a Hypersonic Nozzle*

W. W. WILLMARTH† AND H. T. NAGAMATSU‡
California Institute of Technology, Pasadena, California

(Received April 17, 1952)

A small, stainless steel, two-dimensional source flow nozzle supplied with bottled nitrogen was used for the condensation investigation. It was found that the nitrogen supersaturates approximately 15°K or 1.2 Mach numbers when expanded from stagnation conditions of 70°F and pressures of 8.21 and 16.15 atmos. A numerical method for solving the equations of motion with the aid of the experimental data allows the computation of the fluid temperature during the condensation process. The addition of small amounts of carbon dioxide reduced the degree of supersaturation obtainable with bottled nitrogen.

INTRODUCTION

THE phenomenon of the condensation or sublimation of a gas or vapor when expanded in a Laval nozzle may be simply explained on a qualitative basis. The working fluid undergoes an isentropic expansion with an accompanying decrease in static pressure and temperature. At low temperatures the saturation vapor pressure of the fluid decreases, with temperature, more rapidly than does the actual static pressure in the nozzle, and thus saturated conditions are approached.

This condensation of an expanding vapor was first noted in 1887 by R. von Helmholtz,[1] who also found that steam supersaturated before condensing. The onset and degree of steam condensation were later found, by Stodola[2] and Oswatitsch,[3] to depend on the amount of impurities, in the form of small particles, contained in the steam.

The phenomena described above were similar to the results obtained when humid air was expanded in supersonic wind tunnel nozzles.[3,4] In addition, it was found that when large amounts of water vapor were present in air the collapse of the supersaturated state gave rise to a condensation shock.

When hypersonic wind tunnels were first contemplated, it was realized that the working fluid, usually air, could condense or sublimate (hereafter the single word condensation will be used for simplicity) at the low pressures and temperatures obtained at high Mach numbers. In July, 1950, when this research program was initiated, the fact that air condensed in hypersonic nozzles was generally accepted; and it had been shown, by Becker[5] and Nagamatsu,[6] that condensation occurred, without condensation shock, in the saturation region. Later work, by Stever and Rathbun,[7] established the fact that the condensation of the components of air occurred at modest degrees of supersaturation. These experiments were significant, but they did not indicate the amount of supersaturation obtainable nor the effect of foreign nuclei on the condensation process.

The Volmer theory had been proposed for the prediction of the onset of condensation by Wagner[8] and Charyk and Lees.[9] However, it had not been definitely shown to be capable of predicting the onset of condensation for the components of air. Oswatitsch[3] was able to predict, in a step by step computation, the point in a nozzle expansion at which supersaturated steam first condensed. He also was able to predict the static pressure and amount of condensed steam present after the onset of condensation. In order to bracket the condensation phenomena at the other extreme, the GALCIT group[10] carried out calculations assuming the presence of an infinite number of condensation nuclei so that the condensation of air occurred when saturation conditions were reached.

The present experimental program was initiated with the main purpose of establishing the amount of supersaturation that would be practicably obtainable using nitrogen in a hypersonic nozzle expansion. In addition, experimental information about the fluid properties during the condensation process has been obtained, and the effect of carbon dioxide on supersaturated nitrogen has been investigated.

EXPERIMENTAL EQUIPMENT

The nozzle assembly, piping, and steam heater shown in Figs. 1 and 2 were made of 18–8 stainless steel

* This investigation was sponsored by the Army Ordnance Department and the Air Force.

† Graduate Assistant, Guggenheim Aeronautical Laboratory.

‡ Senior Research Fellow, Guggenheim Aeronautical Laboratory.

[1] R. von Helmholtz, Ann. Physik **32**, 1 (1887).

[2] A. Stodola, *Steam and Gas Turbines* (McGraw-Hill Book Company, Inc., New York, 1927).

[3] K. L. Oswatitsch, "Condensation Phenomena in Supersonic Nozzles" (February, 1942), Z.A.M.M., Vol. 22, No. 1, (German) or R.T.P. Translation No. 1905, (English).

[4] R. M. Head, "Investigation of spontaneous condensation phenomena," California Institute of Technology thesis, (1949).

[5] J. V. Becker, J. Appl. Phys. **21**, 7 (1950).

[6] H. T. Nagamatsu, "Results of recent hypersonic flow research in the Army Ordnance-C.I.T. hypersonic wind tunnel" (May, 1950), Phys. Soc. Meeting, Urbana, Illinois.

[7] H. G. Stever and K. C. Rathbun, "Theoretical and experimental investigation of the condensation of air in hypersonic wind tunnels" (October, 1950), Final Report NACA Contract NAw-5713, or NACA TN 2559.

[8] C. Wagner, "Berechnungen über die Möglichkeiten zur Erziehlung hoher Machscher Zahlen in Windkanaldüsen" (1942), W.V.A., Report No. A-3, Kochel.

[9] J. V. Charyk and L. Lees, "Condensation of the components of air in supersonic wind tunnels" (March, 1948), Princeton Univ. Aero. Engr. Lab., Report No. 127.

[10] R. Buhler, "Recent results of the condensation investigation" (July and August, 1950), Army Ordnance-C.I.T. Hypersonic Wind Tunnel Memoranda.

to prevent corrosion and subsequent contamination of the fluid. The rectangular cross-section nozzle was one inch wide with a maximum test section area of approximately one square inch. The throat, with a radius of 1/16 inch, was adjustable so that the throat area could be varied from 0 to 0.125 square inch. A straight wall nozzle producing an approximation to a source flow was used since, with the proper throat setting, the fluid properties before and after the onset of condensation could be determined in a single test.

Glass side walls were clamped over the lapped nozzle surfaces on which a thin layer of Silicone grease was applied for sealing purposes. The sealing procedure was adequate with the exception of the throat side wall intersection. There, small leaks from the upstream contraction section to the downstream expansion region caused two waves originating near the throat from each side wall. The waves were in a vertical plane and intersected on the tunnel centerline, approximately 1½-inches downstream of the throat. The waves were greatly reduced in magnitude when a rubber-padded C-clamp was tightened on the glass side walls directly over the throat. In no case, with or without clamps, were the waves of sufficient strength to be detected after reflection from the tunnel side walls. In addition, an impact pressure survey to determine the wave angle and Mach number ahead of the waves showed that they were of negligible strength.

The nozzle was supplied with a commercial grade of bottled dry nitrogen (Linde Air Company) which was regulated by an externally loaded regulator (Grove, "Mighty Mite"). The reservoir pressure was controlled by frequent adjustment of the external loading pressure to an accuracy of ½ percent. The reservoir pressure and temperature were measured, respectively, by an Ashcroft Laboratory Test Gauge and a copper-constantan thermocouple mounted in tees just ahead of the nozzle. The supply line, at that point, was standard 1/2-inch stainless steel pipe so that the nitrogen velocity was low, approximately 25 ft/sec. The nozzle was exhausted by two Fuller rotary vane compressors producing an absolute pressure of 1½ cm Hg when the nozzle was in operation.

The nozzle static pressures were measured through 0.040-inch diameter holes staggered along the upper and lower nozzle surfaces 1/2-inch apart. The impact pressures were measured at various stations on the nozzle centerline by a 3/4-inch length of 0.043-inch diameter stainless steel tubing, with an inside diameter of 0.026 inch. This was mounted in a 1/8-inch brass tube which ran through the hole in the wedge at the rear of the nozzle. The wall static pressures were com-

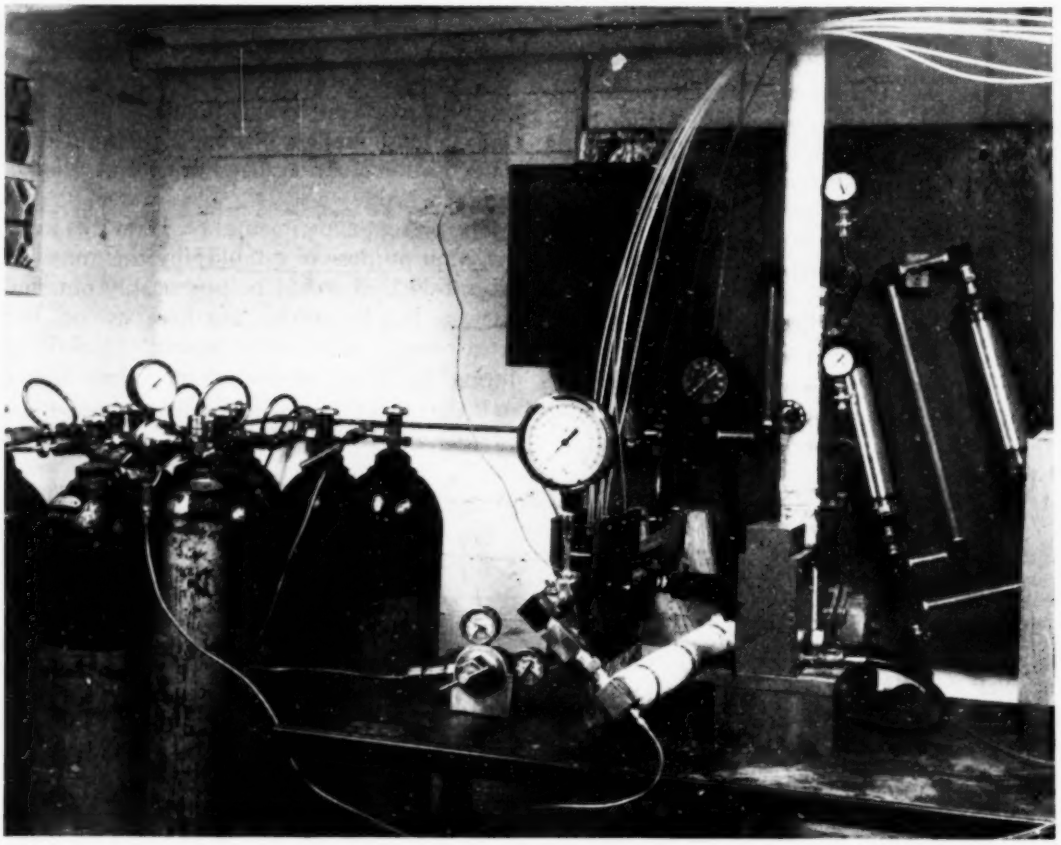


FIG. 1. General view of the experimental equipment.

pared with those measured by a static pressure probe on the nozzle centerline and found to agree within 2 percent at similar nozzle stations. The wall static pressures were measured individually on a bank of silicone manometer tubes referenced to an absolute pressure of 100 microns, whereas the higher impact pressures were measured on a mercury manometer. The accuracy of the pressure measurements was estimated to be within 2 percent for the lowest pressures encountered. A more complete description of the above equipment is available.[11]

FLUID

As mentioned previously, the fluid used for this investigation was bottled dry nitrogen. The choice of nitrogen was influenced by a number of considerations, hence those of primary importance will be enumerated. Nitrogen is diatomic, like air, and is also the main constituent of air. It is not toxic or explosive and is available in reasonably pure form at low cost. In addition, one gas, rather than a mixture, was desirable in that the effect of impurities could be readily determined. Two analyses, by a mass spectrometer, of nitrogen samples taken at the beginning and mid-point of the test program are presented in Table I. Since the mass spectrometer was not sensitive to small amounts of water vapor, a high pressure dew-point meter with a chromium plated indicating button, cooled by carbon dioxide gas, was used to determine the water content of the bottled nitrogen.

RESULTS

The onset of condensation was determined by pressure measurements and a light scattering technique used by other investigators.[5,7] The pressure measurements depended upon the fact that the static pressure in a condensing flow rises above the static pressure predicted by isentropic, perfect gas theory,[5] whereas the impact pressure is almost unaffected by condensation,[11,12] at least for the conditions herein reported. The static pressure rise is due to the heat released by the condensing fluid and should, therefore, represent a criterion for the detection of the onset of condensation. In fact, this is the case as one may verify from the results of a typical run for which both light scattering and pressure measurements were made (see Figs. 3 and 4).

In Fig. 3, the light scattered at an angle of 45° from a 1/4-inch diameter beam of parallel light was recorded at various axial stations in the nozzle. (Five pounds of nitrogen passed through the light beam at each station.)

When no condensation was present, as at the station $X=3.5$ inches, the area between the reflections from the glass sidewalls was black, whereas further downstream

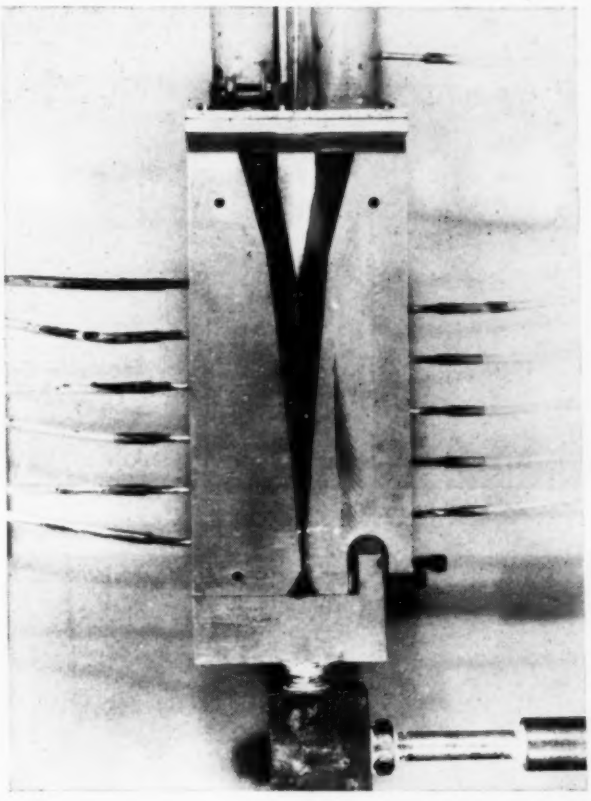


FIG. 2. Side view of the nozzle.

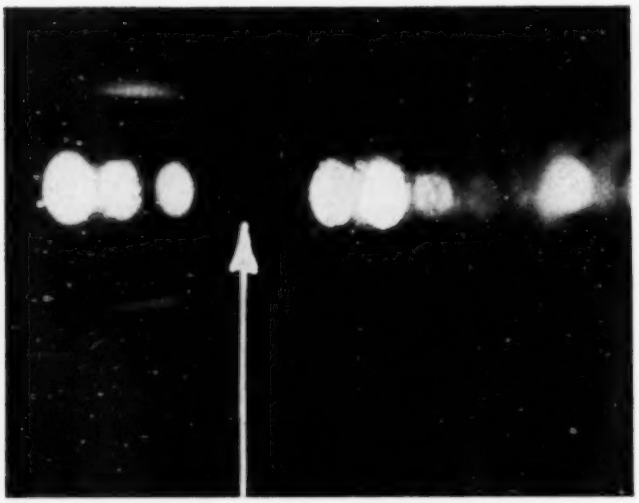
an increasing amount of light scattered from the condensed particles caused a fog to be visible. The pressure measurements (see Fig. 4) are presented as a plot of static pressure divided by impact pressure p/p_0', which is affected by condensation, against the ratio of impact to reservoir pressure p_0'/p_0, which is only slightly affected by condensation. Thus, when condensation occurs the experimental points depart from the isentrope and approach the values predicted for a saturated expansion.[10] From Figs. 3 and 4, Run 9–5, it is apparent that condensation occurred between stations $X=3.5$ and 4.0 inches at a Mach number of approximately

TABLE I. Mass spectrometer analysis of the impurities present in two samples of dry nitrogen taken at the beginning (No. 1) and mid-point (No. 2) of the test program.

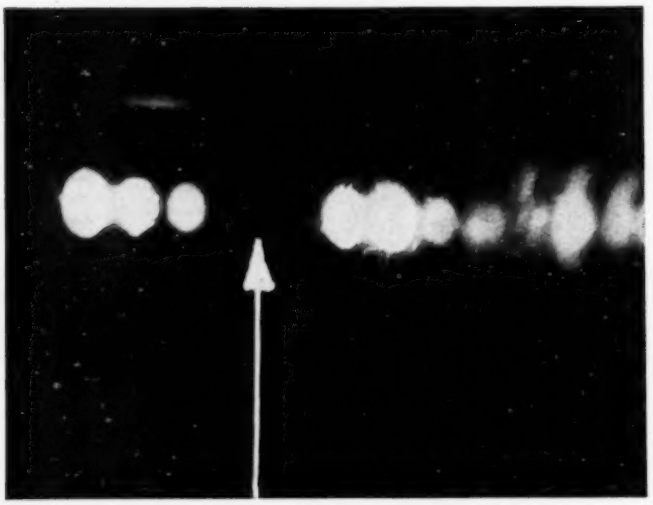
Sample No. 1		Sample No. 2	
Substance	% by volume	Substance	% by volume
Nitrogen	99.81	Nitrogen	99.82
Oxygen	0.12	Oxygen	0.09
Argon	0.07	Argon	0.09
Not found	Detectable	Not found	Detectable
Hydrogen	0.02	Hydrogen	0.004
Helium	0.06	Helium	0.01
Water	0.02	Water	0.003
Carbon dioxide	0.007	Carbon dioxide	0.002
Hydrocarbons	0.004	Hydrocarbons	0.004

[11] H. T. Nagamatsu and W. W. Willmarth, "Condensation of nitrogen in a hypersonic nozzle," GALCIT Hypersonic Wind Tunnel Memorandum No. 6 (January 15, 1952).

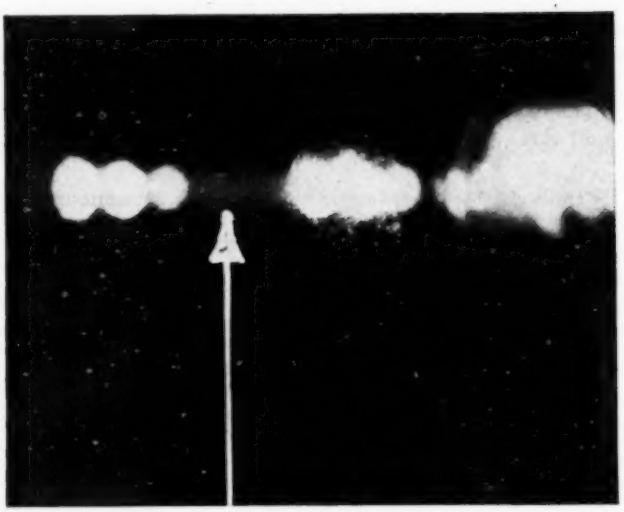
[12] P. Wegener, N.O.L. Hyperballistics Tunnel No. 4, Results I: Air Liquefaction (January, 1951), NAVORD Rep. No. 1742.



a. $X=3.5$ in.



b. $X=4.0$ in.



c. $X=4.5$ in.

FIG. 3. Light scattering pictures used for determining the onset of condensation for Run 9–5, $P_0=8.21$ atmos, $T_0=295°K$, $A^*=0.010$ in.,[3] $M_\infty=4.60$. At the station $X=3.5$ inches from the throat, the Mach number was 5.84 and no fog was recorded. At stations $X=4.0$ and $X=4.5$ inches an increasing amount of fog was recorded.

5.84 and that the static pressure and light scattering measurements correlate with each other.

From the above results, nitrogen condensation was first detected at a Mach number of 5.84, whereas the pressure and temperature at which nitrogen condenses on a plane surface, under equilibrium conditions, were reached at a Mach number, M_∞, of 4.60. Thus, the nitrogen was supersaturated approximately 1.2 Mach numbers. A similar result was obtained with a different reservoir pressure, Run 13–2, Fig. 4.

In order to obtain more information on the state of the nitrogen during the condensation process, a method based on previous work[3,12] has been elaborated. The method will be briefly presented; however, a detailed explanation is available.[11]

The equations of continuity,

$$\rho uA=\text{const}, \tag{1}$$

momentum,

$$udu+dp/\rho=0, \tag{2}$$

energy,§

$$udu+c_p dT-ldg=0, \tag{3}$$

and a modified equation of state,

$$p=(1-g)\rho RT, \tag{4}$$

govern the flow process in a nozzle expansion where the fluid flow may be isentropic, $g=0$, or condensing, $g\neq 0$. Here, g is the mass fraction of fluid condensed, u is the velocity of the vapor and condensed particles, ρ is the density of the mixture, p and T are the static pressure and temperature of the vapor, c_p is the specific heat of the vapor, l is the latent heat, assumed constant, and A is the effective channel area as determined from the impact pressure using the isentropic perfect gas relations. Equations (1) and (2) may be combined, to give an expression for the velocity ratio,

$$u(x)/u^*=(\gamma+1)/\gamma-(1/\gamma)[(\gamma+1)/2]^{\gamma/\gamma-1}\times\left\{[(A(x)/A^*)(p(x)/p_0)]-I-\int_{x_1}^{x}(p(x)/p_0)\frac{d(A(x)/A^*)}{dx}dx\right\}, \tag{5}$$

where

$$I=\int_{x^*}^{x_1}(p(x)/p_0)\frac{d(A(x)/A^*)}{dx}dx=\left[\left(\frac{u(x)}{u^*}\right)_{x_1}-\frac{\gamma+1}{\gamma}\right]\Big/\left[\frac{1}{\gamma}\left(\frac{\gamma+1}{2}\right)^{\gamma/\gamma-1}\right]+\left[\left(\frac{A(x)}{A^*}\right)\left(\frac{p(x)}{p_0}\right)\right]_{x_1}.$$

§ This energy equation is approximate in that the energy of the liquid phase has been neglected. The error involved is small and has been included in the estimate of the accuracy of the computation which is stated below.

Here, X_1 is a station downstream of the throat station, X^*, but ahead of the condensation zone; hence, the velocity ratio and effective area ratio may be determined from pressure measurements using ordinary channel flow equations. Thus, the velocity ratio at any station, X, may be determined from Eq. (5), since the area ratio is given for isentropic or condensing flow by the impact pressure and isentropic, perfect gas relations. It has been shown experimentally,[11] by making heated and unheated runs, that the impact pressure is almost a unique function of the area ratio for isentropic or condensing flow. In addition this result is predicted by saturated expansion theory.[10,11]

Using Eqs. (1) and (4) the temperature ratio is

$$T(x)/T^* = [(p(x)/p^*)(u(x)/u^*)(A(x)/A^*)]/[1-g(x)]. \quad (6)$$

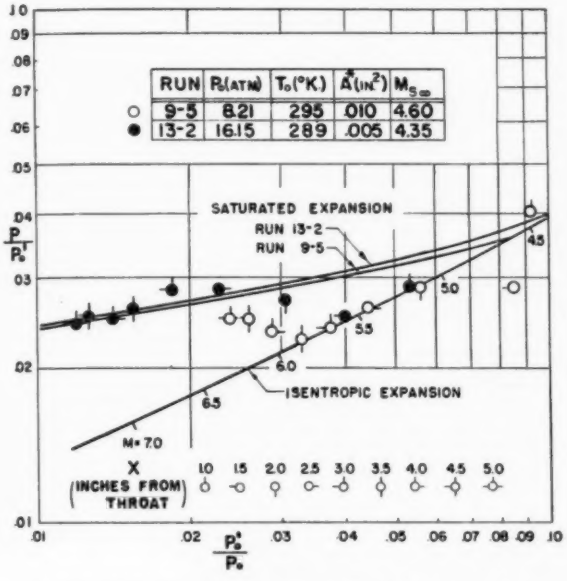


FIG. 4. p/p_0' *versus* p_0'/p_0. The effect of condensation on static pressure.

From Eq. (3), the amount of condensed fluid is

$$g(x)=\left[\frac{\gamma-1}{2}\left(\frac{u(x)}{u^*}\right)^2+\frac{T(x)}{T^*}-\frac{\gamma+1}{2}\right]\Big/\left[\frac{\gamma-1}{\gamma}\,\frac{l}{RT^*}\right], \quad (7)$$

where l/R may be determined by fitting the Clausius-Clapeyron equation to the saturation curve in the region of expected condensation. With the aid of Eq. (5), Eqs. (6) and (7) may be solved for T and g.

The results of the computations, for Runs 9–5 and 13–2, are presented in Figs. 5 and 6. It should be mentioned that an estimate of the errors in the above method[11] indicated that the computed vapor temperature T was accurate to within 2.5°K, whereas the amount of condensed fluid g may have been in error by 25 percent. From Fig. 5 the temperature after the collapse of the supersaturated state parallels the saturation curve for solid nitrogen.[13,14] Since the vapor pressure and temperature, after the collapse, are approximately the equilibrium values predicted for a saturated expansion, Figs. 4 and 5, it will be assumed that the condensed nitrogen was in the solid phase, although no other evidence can be advanced to support this assumption. The displacement of the experimental points in Fig. 5 to the left of the saturation curve may be due to errors in the computed temperature and the saturation

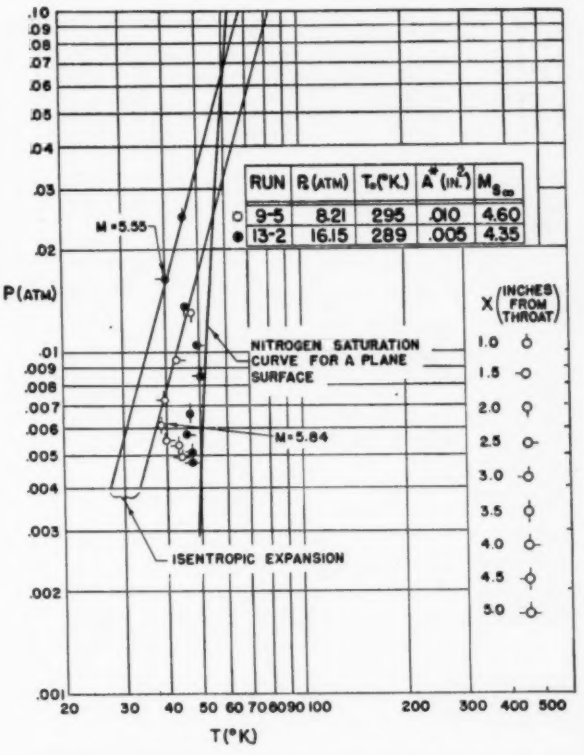

RUN | P_0(ATM) | T_0(°K) | A^*(IN^2) | M_{S_∞}
9-5 | 8.21 | 295 | .010 | 4.60
13-2 | 16.15 | 289 | .005 | 4.35
M=5.55
M=5.84
NITROGEN SATURATION CURVE FOR A PLANE SURFACE
ISENTROPIC EXPANSION
P(ATM)
T(°K)
X (INCHES FROM THROAT)

FIG. 5. Measured pressure *versus* computed temperature during the condensation process.

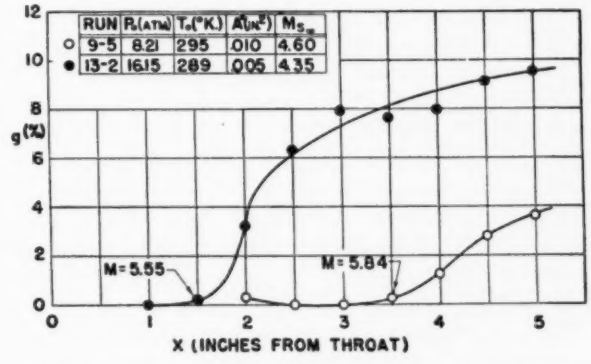


FIG. 6. g *versus* X. The computed amount of condensed phase present during the condensation process.

[13] S. Aoyama and E. Kanda, "The vapour tensions of oxygen and nitrogen in the solid state" (May, 1935), The Science Reports of the Tohoku Imperial University, Sendai, Japan, Vol. XXIV, No. 1.

[14] H. J. Hoge and G. J. King, "The NBS–NACA tables of thermal properties of gases: The vapor pressure of nitrogen" (July, 1950), Table No. 11.50.

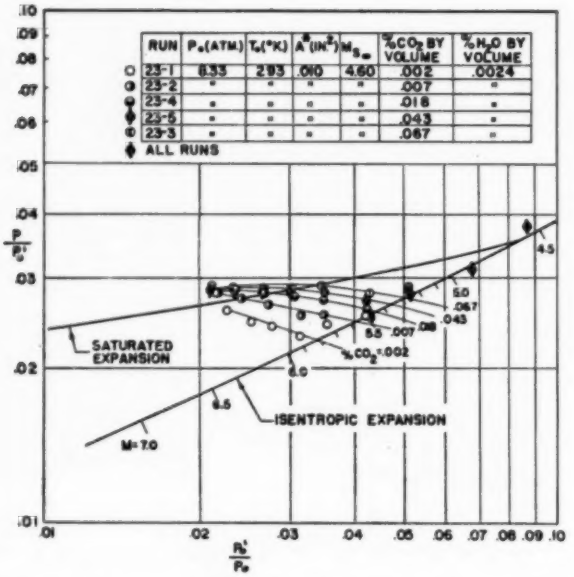


FIG. 7. p/p_0' *versus* p_0'/p_0. The effect of varying amounts of carbon dioxide on the onset of condensation.

curve. Another explanation for the displacement is that one or both of the above errors exist and, in addition, a correction to the vapor pressure accounting for the small size of the condensed particles must be applied.

The time required for the collapse of the supersaturated state may also be estimated from the temperature computation. Using the average computed velocity during collapse, the time required was approximately 5×10^{-5} second for Run 9–5 and 2.5×10^{-5} second for Run 13–2. The increase in area ratio for Run 9–5 was half the increase in area ratio, during collapse, for Run 13–2. Thus, the faster the area ratio increased, the faster the supersaturated state collapsed. However, the amount of supercooling was essentially the same for these two runs, approximately 15°K.

Since the nitrogen became supersaturated during the nozzle expansion, the onset of condensation should occur at a lower Mach number when impurities, acting as condensation nuclei, are added. The effect of adding one impurity, bottled carbon dixode, was investigated. The carbon dioxide was injected through a small orifice eight feet upstream of the nozzle throat. A sample of contaminated nitrogen was withdrawn one foot upstream of the throat and its carbon dioxide concentration was measured using a chemical method developed by Thomas.[15]

The data from four runs with varying carbon dioxide concentration and from a run with no carbon dioxide added are shown in Fig. 7. The lines faired through the experimental points after the onset of condensation show the effect of carbon dioxide concentration on the saturation Mach number, M_s. It can be seen that increasing amounts of carbon dioxide decrease the saturation Mach number. Additional work has shown that larger amounts of carbon dioxide will completely eliminate nitrogen supersaturation.[16] Another effect of carbon dioxide addition was to decrease the impact pressure slightly in the condensation region.[11] The cause of this decrease is not known.

From the data of Fig. 7, the critical supersaturation pressure ratio, p/p_∞, has been obtained and plotted against carbon dioxide concentration (see Fig. 8). In this plot, p_∞ is the vapor pressure of a plane surface of nitrogen at the temperature of the fluid when condensation began, and p is the actual pressure of the supersaturated vapor at the onset of condensation. The onset of condensation was assumed to be the intersection of the faired lines and the isentrope of Fig. 7. The dark circle plotted on Fig. 8 was obtained from the report of Charyk and Lees[9] and shows the value of p/p_∞ predicted by the Volmer theory for condensation into the liquid phase with $\ln J=7$. This low value of p/p_∞, predicted by the Volmer theory, will not have any significance if the nitrogen was actually in the solid state. Since data on the surface free energy of nitrogen could not be found, no statement as to the applicability of the Volmer theory can be made.

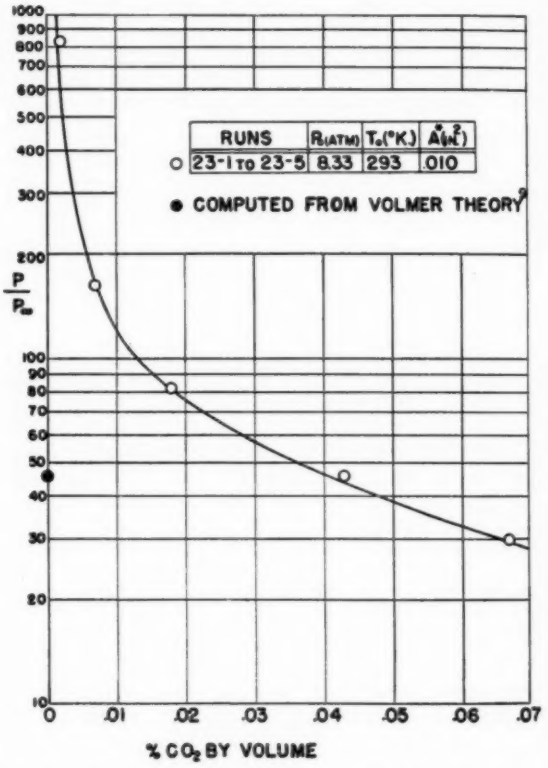


FIG. 8. p/p_∞ *versus* percent carbon dioxide. The effect of varying amounts of carbon dioxide on the critical supersaturation pressure ratio.

[15] M. D. Thomas, "Precise automatic apparatus for continuous determination of carbon dioxide in air" (May, 1933), Industrial and Engineering Chemistry, Analytical Edition, Vol. 5, No. 3.

[16] P. D. Arthur and H. T. Nagamatsu, "Effects of impurities on the supersaturation of nitrogen in a hypersonic nozzle," GALCIT Hypersonic Wind Tunnel Memorandum No. 7, March 1, 1952.

CONCLUSIONS

The light scattering and pressure measurements used to detect the onset of condensation were found to agree within the experimental accuracy of the measurements. The method based on a plot of the pressure ratios p/p_0' and p_0'/p_0 was considered the most desirable due to its simplicity of application.

Commercial bottled nitrogen becomes supersaturated approximately 1.2 Mach numbers when expanded in a straight wedge nozzle. This degree of supersaturation applies only to expansion from reservoir conditions of approximately 70°F and 8 to 16 atmos. With the above reservoir conditions, the nitrogen condensed at temperatures and pressures in the range where solid nitrogen ordinarily exists under equilibrium conditions.

A method for computing the free stream temperature of the saturated vapor using the measured impact and static pressures has been elaborated. The method assumes that reversible phase changes occur and, within the assumption of reversibility, the temperature computation has been shown to be accurate to within 5 percent.

An attempt to further purify the nitrogen by removing carbon dioxide, water vapor, and small particles had no appreciable effect on the amount of supersaturation obtainable, although the reaction time of the chemicals was varied by deepening the chemical bed. The nitrogen may have condensed due to self-nucleation, although measurable amounts of oxygen and argon were present. It is not known whether the degree of supersaturation would have been increased by removing the oxygen and argon.

Various amounts of carbon dioxide were added to the nitrogen and were found to decrease the critical supersaturation ratio and the Mach number at which the nitrogen condensed. The largest amount of carbon dioxide added, 0.067 percent by volume, did not decrease the condensation Mach number to the value predicted by the saturated vapor data for a plane surface of nitrogen. Addition of larger amounts of carbon dioxide did eliminate the supersaturation of nitrogen as found by later investigation.[16] The impact pressures measured downstream of the onset of condensation were found to be decreased by the addition of carbon dioxide to the flow. The cause of this decrease is not known.

Two runs were made with different expansion ratios. The degree of supersaturation was independent of the time required for expansion when the expansion ratio was changed by a factor of two.

The saturated expansion theory[10] closely approximates the fluid properties after the onset of condensation, even when the nitrogen supersaturates before the onset of condensation and then, presumably, condenses into the solid phase.

Announcement

The tenth annual meeting of the Electron Microscope Society of America will be held at the Statler Hotel in Cleveland, Ohio, Thursday through Saturday, November 6, 7, and 8, 1952. Programs and abstracts of papers to be presented will be available about October 15 from Dr. Carl E. Willoughby, Program Chairman, du Pont Experimental Station, Wilmington 98, Delaware. Information about local arrangements may be had from Mrs. Mary S. Jaffe, LDL-85, General Electric Company, Nela Park, Cleveland 12, Ohio.

Influence of Initial Velocities on Electron Transit Time in Diodes

J. T. WALLMARK
Royal Institute of Technology, Stockholm, Sweden
(Received March 17, 1952)

The transit time spread due to difference in initial velocity in a planar diode has been calculated for the case of nonhomogeneous velocity distribution. The spread is given as a function of the current for two practical cases, the normal diode and an inverted diode (a beam reflected in a diode space). The spread is also given as a function of anode voltage within certain limits. The method may be used to estimate the influence of transit time spread on transit time devices. A special application is described.

ELECTRON transit times in electronic apparatus are of importance whenever the frequencies used are so high that the transit time becomes an appreciable fraction of the cycle. The spread in transit time is usually of little concern. However, in electronic apparatus where the transit time itself is directly utilized, such as reflex klystrons or Barkhausen-Kurz oscillators, the spread in transit time plays an important role.

Calculations on spread in transit time caused by differences in initial velocities have been reported by Cockburn[1] and Barut,[2] who assume a homogeneous velocity distribution. In practical cases calculations for a nonhomogeneous velocity distribution are of more interest. The author has undertaken such calculations by means of a first-order perturbation method.

It has been well established that the velocities of electrons originating from a thermionic cathode have a Maxwellian distribution as in Fig. 1. Practical electronic devices are with few exceptions run with space charge limited current to take advantage of the stable conditions thus obtained. The velocity distribution, however, is still Maxwellian, as the presence of a virtual cathode just outside the cathode surface has the same effect on the current as a change in work function and does not affect the distribution.

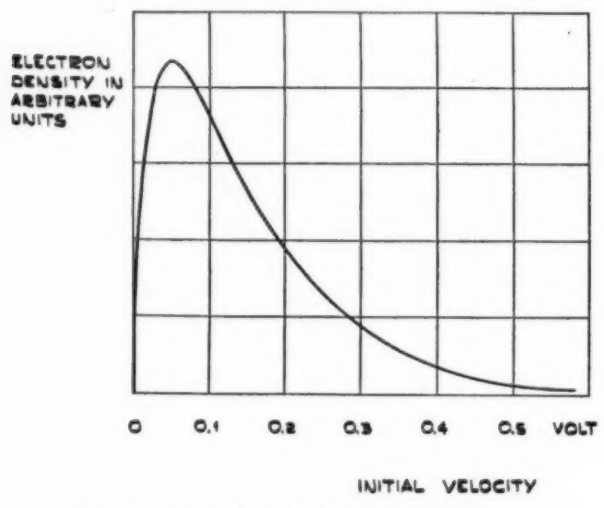


FIG. 1. Distribution of initial velocities from an oxide cathode at 1160K°.

METHOD

Then a first-order approximation can be made assuming a potential distribution corresponding to that case in which nearly all of the electrons are emitted with the initial velocity δ. The transit time for a number of electrons emitted with a velocity that differs from the main velocity by the amount Δ can then be calculated if we assume that the number is so small that the electric field is not disturbed. For the velocity δ was chosen the average velocity 0.1 volt which is characteristic of an oxide emitter at 1160°K. In a preliminary publication[3] the maximum of the distribution curve was chosen with nearly the same results.

To calculate the spread in transit time let us neglect the few electrons with very high and very low initial velocity and consider only the main part. Let us as arbitrary boundaries for this main part take the half-values of the distribution function of Fig. 1. This gives as boundaries 0.0135 and 0.185 v. The spread in transit time calculated for these boundaries has to be taken as a minimum value, as there are still 30 percent of the electrons with wider spread.

CALCULATIONS

Assume two infinite parallel planes with an electron flow perpendicular to them as in Fig. 2. The electric field strength, potential, etc., may be calculated from published data on partial space charge.[1,2,4-6] Consider a small number of electrons with a velocity that differs from the main velocity by the amount Δ. The transit time for those electrons is

$$\tau=\int_0^d \frac{dx}{v}.$$

By introducing

$$v=[(V+\Delta)2e/m]^{\frac{1}{2}}$$

$$Edx=dV\Big/\left[\frac{16}{9}\bar{I}\left(\frac{V}{V_A}\right)^{\frac{1}{2}}+\bar{E}_c^2\right]^{\frac{1}{2}},$$

[1] R. Cockburn, Proc. Phys. Soc. **47**, 810 (1935).
[2] A. O. Barut, Z. Angew. Math. u. Phys. **II:1**, 35 (1951).
[3] J. T. Wallmark, Phys. Rev. **84**, 598 (1951).
[4] H. F. Ivey, Phys. Rev. **76**, 554 (1949).
[5] J. T. Wallmark, Proc. Inst. Radio Engrs. **40**, 41 (1952).
[6] W. M. Brubaker, Phys. Rev. **83**, 268 (1951).

where $E=V_A/d$ is the electric field strength without space charge, $\bar{E}_c E$ the electric field strength at the cathode, and $\bar{I}=I/I_{max}$ the relative current where I_{max} is the space charge limited current, dashed letters representing relative values, we obtain

$$\tau=\left(\frac{m}{2e}\right)^{\frac{1}{2}}\int_0^{V_A} dV \Big/ \left\{E\left[\frac{16}{9}\bar{I}\left(\frac{V+\delta}{V_A+\delta}\right)^{\frac{1}{2}} + \bar{E}_c^2\right]^{\frac{1}{2}}(V+\delta+\Delta)^{\frac{1}{2}}\right\}$$

$$\approx\left(\frac{m}{2e}\right)^{\frac{1}{2}}\int_\delta^{V_A} dV \Big/ \left\{E\left[\frac{16}{9}\bar{I}\left(\frac{V}{V_A}\right)^{\frac{1}{2}} + \bar{E}_c^2\right]^{\frac{1}{2}}(V+\Delta)^{\frac{1}{2}}\right\}. \quad (1)$$

The integral is elliptic and an exact solution leads to fairly complicated expressions. However, an approximate solution with satisfactory accuracy can be obtained by a series expansion of $(1+\Delta/V)^{-\frac{1}{2}}$ in the interval $\Delta-V_A$, and $(1+V/\Delta)^{-\frac{1}{2}}$ in the interval $0-\Delta$, and integration of the resulting expression term by term. The series are uniformly convergent within the interval of integration, but not at the boundary Δ. However, the integration can be carried as close to Δ as desired, say to $\Delta\pm\epsilon$, where ϵ is a voltage small compared to Δ. After integration $\Delta\pm\epsilon$ is also expanded into a series and ϵ is made to approach zero.

This general method to calculate the transit time spread cannot be expected to be more accurate than ± 10 percent with the simplifications made. Consequently, terms in the integral (1) which influence the final result less than 10 percent may be neglected.

We obtain for the interval $0-\Delta$

$$\tau_1=C\left\{\frac{7}{8}\left(\frac{\Delta}{V_A}\right)^{\frac{1}{2}}-\frac{2}{7}k\frac{\Delta}{V_A}+\frac{1}{8}k^2\left(\frac{\Delta}{V_A}\right)^{\frac{3}{2}}\right\} \quad (2)$$

and for the interval $\Delta-V_A$

$$\tau_2=\frac{4C}{k}\Bigg|_{1/[k(\Delta/V_A)^{\frac{1}{2}}+1]^{\frac{1}{2}}}^{1/(k+1)^{\frac{1}{2}}}\left\{\frac{1}{x}-A(1+6A)\times\left(\frac{2x}{1-x^2}-\log\frac{1+x}{1-x}\right)-A^2x^3\frac{9x^2-5}{(1-x^2)^3}\right\}, \quad (3)$$

where

$$C=d/[\bar{E}_c(V_A 2e/m)^{\frac{1}{2}}], \quad A=\Delta k^2/8V_A,$$
$$X=1/[k(V/V_A)^{\frac{1}{2}}+1]^{\frac{1}{2}}, \text{ and } k=16\bar{I}/9\bar{E}_c^2.$$

The method outlined has been used to obtain the spread in transit time for the two practical cases following.

PLANAR DIODE

Figure 3 shows the conditions for a planar diode with the data $d=5$ mm, $V_A=100$ volts. The diagram gives transit time as a function of initial velocity with the current as parameter. Both transit time and spread in transit time for $\bar{I}=1$ are calculated for the distance between the virtual cathode and the anode, and the additional transit time within the virtual cathode has not been dealt with. Figure 5 gives the spread in transit time as defined above. As can be seen the spread in transit time increases with current and amounts to approximately 7 percent at space charge limitation. For comparison also the distribution curve of Fig. 1 has been redrawn in connection with Fig. 3.

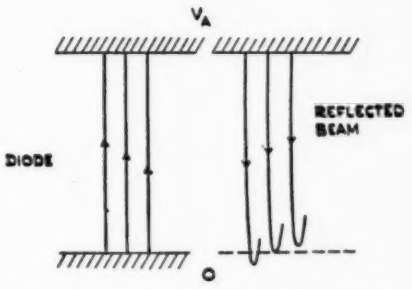


FIG. 2. The two cases considered; a diode and an inverted diode (reflected beam).

REFLECTED BEAM

Figure 4 shows the half-transit time for a reflected beam in a planar diode space with the same data as above and as shown in Fig. 2. The electric field beyond the turning point for the main part of the electrons has

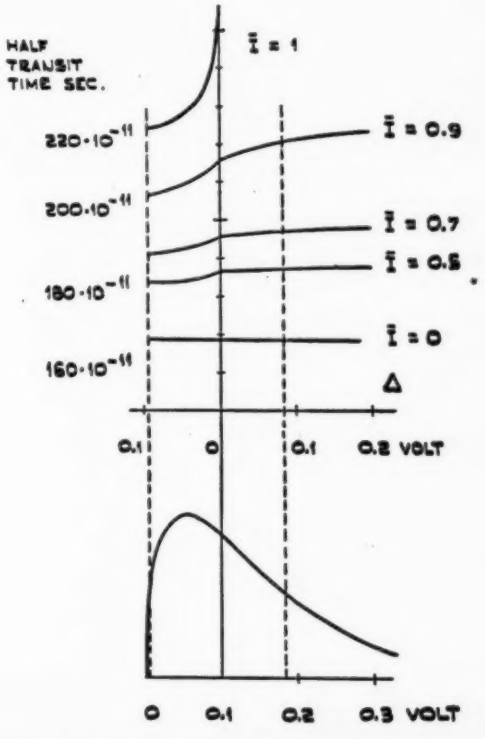


FIG. 3. Transit time *vs* initial velocity with current as parameter in a planar diode with $d=5$ mm, $V_a=100$ volts.

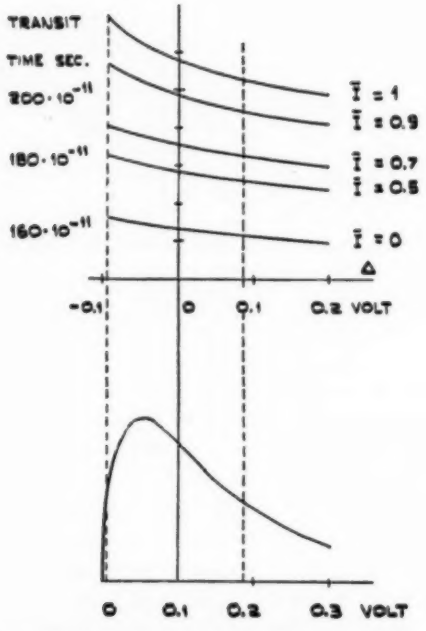


FIG. 4. Half-transit time *vs* initial velocity with current as parameter in an inverted diode with $d=5$ mm, $V_a=100$ volts.

been assumed constant and equal to the field at that point.

The transit time spread at low currents is vanishingly small. At $\bar{I}=1$, space charge limitation, and for positive values of Δ the method indicates infinite spread. This would mean that electrons with high velocity should penetrate to the other side of the turning point for the main part of the electrons, where the electric field strength is zero, and get lost. In the practical case this is not possible. However, an approximate value of the transit time spread has been obtained by assuming $\delta=0.185$ volt and calculating the spread for negative values of Δ. This gives a spread of 13.5 percent. The results are shown in Figs. 4 and 5.

It is to be noted that a high velocity gives a long transit time in contrast to the previous case of a planar diode where a high velocity instead gives a short transit time.

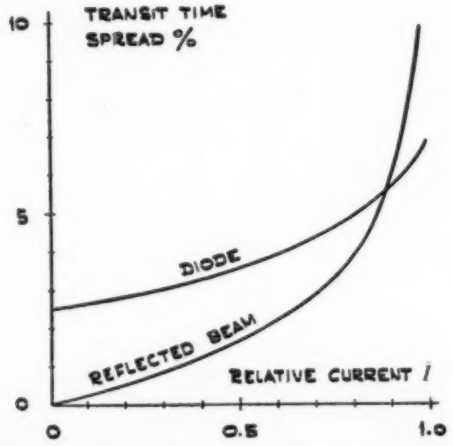


FIG. 5. Transit time spread in a diode and in an inverted diode.

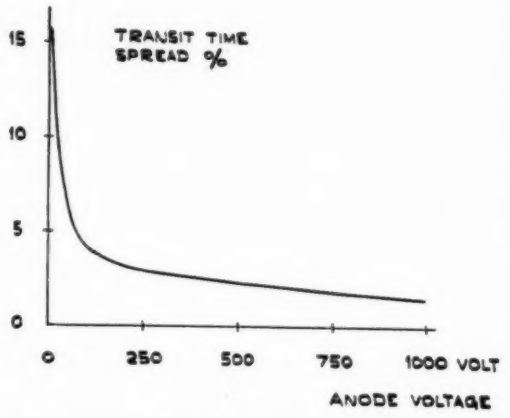


FIG. 6. Influence of anode voltage on spread in transit time for a diode with $d=5$ mm.

Although the foregoing results are valid for fixed values of anode voltage and cathode-anode distance, it is of interest to study the influence of a variation in these parameters. From the expressions (2) and (3) it is clear that an increase of V_A or a reduction of d reduces the spread in transit time. For comparison the spread has been calculated in the planar diode case for anode voltages from 10 to 1000 volts as shown in Fig. 6, but for constant space charge conditions ($\bar{I}=0.7$). If instead the current is kept constant the curve becomes much steeper.

A SPECIAL APPLICATION

The results obtained here have been used to estimate the influence of initial velocities in the special case of a new electron tube utilizing space charge beam deflection[4] and characterized by a very high transconductance-to-current ratio. The principle may be stated the following way.

Consider a beam of electrons ejected at an arbitrary angle from a plane at positive potential towards another plane parallel to the first and at a negative potential.

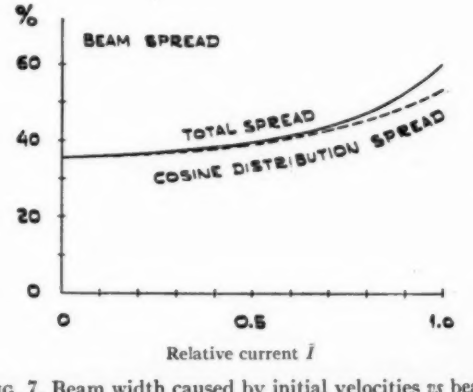


FIG. 7. Beam width caused by initial velocities *vs* beam current for a reflected beam.

The beam is reflected and returns to the first plane in a parabolic path. Increased beam current and consequently increased space charge increase the transit time. This results in a wider arc, or in other words, a longer traveled distance parallel to the planes, as the velocity parallel to the planes is not changed. The increase in traveled distance parallel to the planes is converted to an additional transconductance in the tube by arranging the anode so that the percentage of beam current reaching the anode varies as the beam current.

However, initial velocities of the electrons cause differences in transit time and, consequently, spreading of the beam. As the transconductance of the tube is proportional to the current density of the beam, it is reduced by the spreading.

At the same time spreading is introduced by the cosine distribution in space of the initial velocities. This spreading is directly proportional to the transit time and rises from 36 percent at $\bar{I}=0$ to 54 percent at $\bar{I}=1$ with the same data on the reflection space.

The resultant influence of initial velocities on the spread is shown by Fig. 7.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 10 OCTOBER, 1952

On a Multiple Scattering Theory of the Finite Grating and the Wood Anomalies*,†

V. TWERSKY
Mathematics Research Group, New York University, Washington Square College, New York, New York

(Received April 10, 1952)

The formal multiple scattering solution obtained previously for the scattering of a plane wave by an arbitrary configuration of parallel cylinders is applied to the finite grating of cylinders and to the reflection grating of semicylindrical bosses on a perfectly conducting infinite plane. A far field solution in closed form for spacing large compared to wavelength and radius is obtained on neglecting the "end effects." It is found that for the transmission grating both polarizations may be markedly affected by multiple scattering; for the reflection grating, however, the effects are much more pronounced for polarization perpendicular to the elements. The case of radii$\ll\lambda$ is investigated in detail; the λ's for which the effects are greatest are determined, and various intensity curves are presented. The bright and dark bands appearing on these curves (overlaying the usual continuous spectra observed with broadband radiation) are similar to the "grating anomalies" discovered by Wood in 1902. The present analysis provides a simple physical interpretation of their presence in terms of the magnitudes and phases of the various orders of scattering. Simple relations for their dependence on the parameters are derived which agree in general with experimental results. Both the electromagnetic and acoustic cases are treated.

I. INTRODUCTION

THE scattering of a plane wave by a finite grating of parallel circular cylinders of infinite length was treated by Schaefer and Reiche[1] by means of the single scattering approximation, i.e., they neglected excitations of each cylinder arising from the waves scattered by the remaining cylinders. To obtain a criterion for this approximation, the writer considered each cylinder as excited by the incident plane wave plus the waves scattered by its nearest neighbors as a result of plane wave excitation.[2] It was shown that each term arising from these additional excitations approached zero more rapidly with decreasing radius or increasing spacing than did the single scattered terms. The rapid increase of the number of terms as more neighbors and more involved multiple scattering processes were taken into account, however, indicated that a more thorough investigation should be undertaken. In this paper the scattering of electromagnetic and acoustic waves by the finite grating will be treated by specializing the formal multiple scattering solution we have since obtained for an arbitrary configuration of parallel cylinders.[3,4]

In Sec. II, the complete expression for the scattered wave is given and specialized to grating width small compared to the distance of observation and spacing large compared to wavelength—the case of practical interest. The "end effects" are then neglected to obtain a closed form representation involving only the known solution for the single cylinder and the sum analogs of the Fresnel integrals. Approximate solutions for radii very small compared to wavelength are obtained and investigated in detail in Sec. III. In Sec. IV we consider the analogous grating of semicylindrical bosses on

* This work was performed at Washington Square College of Arts and Science, New York University and was supported in part by Contract No. AF-19(122)-42, with the U. S. Air Force through sponsorship of Geophysics Research Division, Air Force Cambridge Research Center, Air Materiel Command.

† This material is part of the writer's "Multiple Scattering of Radiation, Part II," Research Report EM-39, Mathematics Research Group, New York University, Washington Square College, 1951.

[1] C. Schaefer and F. Reiche, Ann. Physik **35**, 817 (1911).

[2] V. Twersky, J. Acoust. Soc. Am. **22**, 539 (1950).

[3] V. Twersky, J. Acoust. Soc. Am. **24**, 42 (1952). In this paper a general expression for an arbitrary configuration, satisfying any of the usual prescribed boundary conditions simultaneously at each cylinder, was obtained by explicitly taking into account all multiple excitations of each cylinder arising from the radiation scattered by the others. The scattered wave was expressed as the sum of an infinite number of "orders of scattering," the first being the usual single scattering approximation.

[4] V. Twersky, J. Appl. Phys. **23**, 407 (1952). In Sec. II of this paper the result of reference 3 was specialized for an arbitrary planar configuration, and in Secs. III–V, scattering by two cylinders was treated in detail. The sections and equations of this paper will be indicated by the superscript 4 in the text, this material being referred to wherever feasible to avoid repetition.

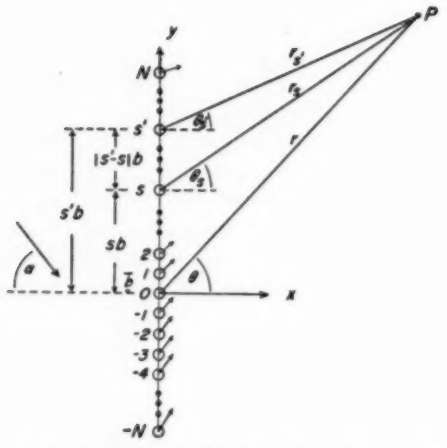


FIG. 1. Plane wave incident on a finite grating of $\mathfrak{N}=2N+1$ parallel cylinders.

a perfectly conducting (or rigid) infinite plane, and in Sec. V the extension of the present theory to gratings with elements other than cylinders is discussed.

Employing the approximate solutions of Sec. III we obtain an analytic criterion for the range of validity of the single scattering approximation and determine the wavelengths for which the effects of multiple scattering are "maximal" (i.e., give rise to the greatest departures from the predictions of single scattering theory). It is of interest to note that the predicted maximal effects are in many respects similar to the "grating anomalies" or "Wood anomalies" observed by Wood,[5] Strong,[6] Ingersoll[7] and Palmer[8] in the white light spectra of certain reflection gratings. The present analysis, based on the magnitudes and phases of the various orders of scattering, provides a simple and physically plausible interpretation of the pronounced bright and dark bands observed for certain wavelengths; this analysis also correlates more of the observed phenomena than does Rayleigh's "dynamical theory of the grating."[9] Frequent references to and comparisons with these experimental results will be found in the following sections.

Because of the complexity of the solution, an intuitive analysis of the intensity-wavelength curve, based on an analog of the vibration curves of diffraction optics, is provided to supplement the results of numerical calculations. The "vibration curve" is also used to derive simple, if approximate, relations for such measurables as the positions of the peaks of the bands, the ratio of band widths for different wavelengths, their resolution, etc.

II. THE FINITE GRATING OF PARALLEL CYLINDERS

Complete Solution

Consider a plane wave $\psi_i=\exp[iKr\cos(\theta+\alpha)]$—the velocity potential for acoustics or the scalars of the Hertz vector potentials for electromagnetics (see Appendix[4])—incident on a grating of $\mathfrak{N}=2N+1$ cylinders of radii a and spacing b as in Fig. 1. For symmetry purposes the reference origin of coordinates r, θ is taken on the axis of the central cylinder at its intersection with the xy plane.

The scattered wave is obtained by specializing our previous result for an arbitrary planar configuration (3)[4] to the case where the cylinders are identical, the s'th cylinder lies a distance sb from the reference origin, and the separation of any two cylinders is $|s'-s|b$. Thus we obtain

$$\psi=\sum_{m=1}^{\infty}\psi^m=\sum_{s=-N}^{N}\sum_{m}{}^s\psi^m$$

$$=\sum_{n=-\infty}^{\infty}A_n i^n\sum_s H_n(Kr_s)e^{in\theta_s-isA}\sum_m {}^sB_n{}^m,$$

$${}^sB_n{}^m=\left[\prod_{\mu=(1)}^{(m-1)}\sum_{n\mu}A_{n\mu}S_{n\mu-1n\mu}\right]\exp[in^{(m-1)}\alpha],$$

$$S_{n\mu-1n\mu}=\exp[i(n^\mu-n^{\mu-1})\pi/2]$$
$$\times\sum_{s^\mu}{}'H_{n\mu-1-n\mu}(|s^\mu-s^{\mu-1}|Kb)$$
$$\times\exp\{i\epsilon_{s^\mu-1s^\mu}[|s^\mu-s^{\mu-1}|A+(n^{\mu-1}-n^\mu)\pi/2]\},\quad(1)$$

where the A_n are the scattering coefficients of the single cylinder (e.g., $A_n=-J_n(Ka)/H_n(Ka)$ for $\psi+\psi_i$ to vanish at each cylinder, etc.), the H's are the Hankel functions of the first kind corresponding to outgoing waves for the suppressed $\exp(-i\omega t)$, and r_s and θ_s locate an observation point P in the coordinate frame of the s'th cylinder; $A=Kb\sin\alpha$, $\sum'$ indicates that $s^\mu=s^{\mu-1}$ is to be excluded from the sum, and $\epsilon_{s^\mu-1s^\mu}=\mp1$ for $s^\mu\gtrless s^{\mu-1}$. For $m=1$, the first order of scattering, the formal product over μ (μ indicating the number of primes distinguishing the indices) reduces to unity, and the sum over m (m indicating the orders of scattering) reduces to $\exp(in\alpha)$. This expression when added to ψ_i constitutes a formal solution for the problem of the grating.

We split the sums over s^μ into two partial sums corresponding to $s^\mu>s^{\mu-1}$, $s^\mu<s^{\mu-1}$, and introduce the positive separation index $\sigma^{\mu-1}=|s^\mu-s^{\mu-1}|$ to obtain

$$S_{n\mu-1n\mu}=\exp[i(n^{\mu-1}-n^\mu)\pi]$$
$$\times\sum_{\sigma^{\mu-1}=1}^{N-s^\mu-1}H_{n\mu-1-n\mu}(\sigma^{\mu-1})\exp(-i\sigma^{\mu-1}A)$$
$$+\sum_{\sigma^{\mu-1}=1}^{N+s^\mu-1}H_{n\mu-1-n\mu}(\sigma^{\mu-1})\exp(i\sigma^{\mu-1}A),\quad(2)$$

[5] R. W. Wood, Phil. Mag. **4**, 396 (1902); **23**, 310 (1912); Phys. Rev. **48**, 928 (1935).
[6] J. Strong, Phys. Rev. **49**, 291 (1936).
[7] L. R. Ingersoll, Astrophys. J. **51**, 129 (1920); Phys. Rev. **17**, 493 (1921).
[8] C. H. Palmer, Doctoral Dissertation, The Johns Hopkins University, 1951.
[9] Rayleigh, Proc. Roy. Soc. London **A79**, 399 (1907); Phil. Mag. **14**, 60 (1907); see also W. Voigt, Göttinger Nachr. 41 (1910); Nachr. Math. Phys. **K1**, 40 (1911). U. Fano, Phys. Rev. **50**, 573 (1936); **51**, 288 (1937); Ann. Physik (5) **32**, 393 (1938). K. Artmann, Z. Physik **119**, 529 (1942). These writers analyze reflection from a corrugated surface whose cross section is representable by a Fourier series. Artmann also presents an intuitive interpretation based on Huygens' principle and multiple scattering—an approach first suggested by Wood (see reference 5, 1935).

where Kb has been suppressed for brevity, and where $s^{\mu-1}\neq N$ in the first sum and $s^{\mu-1}\neq -N$ in the second. The quantity $S_{n\mu-1n\mu}$, the "multiple excitation factor," consists essentially of two finite Schlömilch series for which, unfortunately, no explicit representations exist. The present problem is therefore much more complicated than the two cylinder problem[4] which involved only the first term of a series of this type.

It can be seen that the first order of scattering comprises $2N+1$ terms corresponding to the waves singly scattered by each cylinder; the second order comprises $2N(2N+1)$ terms, each of the $2N+1$ cylinders scattering $2N$ waves as a result of excitation by the waves scattered by the remaining $2N$ elements; and similarly the m'th order comprises $(2N)^{m-1}(2N+1)$ terms. Differences arise in the upper limits of the corresponding sums over s and σ^μ of the various orders as a consequence of the finite number of elements, i.e., of the "end effects" of the finite grating. Thus the exclusion of a term $s=N$ from ψ^2 merely indicates that the N'th cylinder has none above it to contribute to its excitation. Similarly in ψ^3, certain terms for $s=N-1$ are excluded, since $s=N-2$ is the highest cylinder whose tertiary excitations arise from the second order of scattering of a cylinder lying above it, which in turn possesses a secondary excitation from a still higher cylinder.

Far Field Solution

We proceed as Schaefer and Reiche[1] to the far field for distance of observation large compared to grating width. For $Kr\gg 1$, $Kr\gg|n|$ (the second condition implying that the A_n decrease rapidly with increasing n), we can employ $H_n(Kr_s)\rightarrow(2/i\pi Kr_s)^{\frac{1}{2}}i^{-n}\exp(iKr_s)=H_0(r_s)i^{-n}$, where we have suppressed K and where it is understood that H_0 is the asymptotic form. In addition, for $r\gg\mathfrak{N}b$ we have $r_s\approx r-sb\sin\theta\equiv r-sT$ and $\theta_s\approx\theta-(sb/r_s)\cos\theta\rightarrow\theta$, so that $H_0(r_s)\rightarrow H_0(r)\exp(-isT)$ when we neglect $sT\ll r_s$ in the denominator of the square root in $H_0(r_s)$. Hence (1) reduces to

$$\psi=H_0(r)\sum_n A_n e^{in\theta}\sum_s e^{-is\gamma}\sum_m {}^sB_n{}^m,$$
$$\gamma=T+A=Kb\,(\sin\theta+\sin\alpha). \quad (3)$$

Thus the first order of scattering is

$$\psi^1=H_0(r)\sum_n A_n e^{in(\theta+\alpha)}\sum_s e^{-is\gamma}\equiv H_0C\Delta(\mathfrak{N}), \quad (4)$$

where the argument of H_0 has been suppressed, C is the "single cylinder factor" $\sum\epsilon_n A_n\cos n(\theta+\alpha)$ [$n=0$ to $n^*\approx x+2$; $x=Ka$ or $|\nu|Ka$, where ν is the index of refraction; $\epsilon_n=1$ for $n=0$, $\epsilon_n=2$ for $n>0$], and

$$\Delta(\mathfrak{N})=\sum e^{-is\gamma}=\sin(\mathfrak{N}\gamma/2)/\sin(\gamma/2); \quad (4')$$

$\Delta(\mathfrak{N})$ is the usual Fraunhofer "grating factor." To this order of scattering the solution is of the form obtained in elementary treatments of a finite grating of slits, the far field solution of the single cylinder taking the place of the "single slit factor" as the "envelope" of the rapidly oscillating $\Delta(\mathfrak{N})$. Equation (4) for normal incidence is treated in detail in reference 1.

The principal maxima of ψ^1 (neglecting the envelope) occur when

$$\gamma=2q\pi, \quad q=0,\pm1,\pm2,\cdots, \quad (5)$$

for which values $\Delta(\mathfrak{N})$ reduces to $\mathfrak{N}$, the q's being the spectral orders of the grating. Physically these occur when the path difference of the primary excitation at adjacent cylinders plus their path differences to the observation point equals an integral number of wavelengths. The minima of $\Delta(\mathfrak{N})$, equaling zero, occur when $\mathfrak{N}\gamma=2q\pi$, $\gamma\neq 2q\pi$; hence there are $\mathfrak{N}-1$ minima and $\mathfrak{N}-2$ small secondary maxima between two principal maxima. We employ the notion of a "Δ-function in $\mathfrak{N}$" to indicate this behavior, and will concern ourselves only with the principal maxima in following discussions. When white light or other broadband radiation is incident on a grating governed by $\Delta(\mathfrak{N})$, the central maximum at $\theta=-\alpha$ (or at $\theta=\pi+\alpha$) appears white, while on each side appear the various orders of continuous spectra. Although the situation differs if (3) obtains, we still accept $\gamma=2q\pi$ as a labeling device to associate a particular wavelength with a set of angles of observation (or an interrelated sequence of wavelengths with a particular angle). We consider the terms in γ appearing in all ψ^m as yielding the angular distribution of the total scattered radiation, the remaining factors being essentially only source strengths whose values can modify the magnitude of a particular angular component but not alter its position.

Spacing Larger than Wavelength

It is evident from (5) that the spacing must be large compared to wavelength for several spectral orders to appear; we now introduce this restriction explicitly. For $Kb\gg 1$, $Kb\gg|n^{\mu-1}-n^\mu|$ (or essentially $b\gg a$, λ so that each element is in the far field of all the others), we employ $H_{n-n'}(\sigma)\rightarrow H(\sigma)i^{n'-n}$, etc., where $H(\sigma)\equiv h(\sigma)\exp(i\sigma)$ is the asymptotic form of $H_0(\sigma Kb)$. On substituting into (3) and summing the terms for positive and negative values of n^μ, we obtain

$$\psi=H_0\sum_s e^{-is\gamma}[C+Z(s)],$$
$$Z(s)=\sum_{n,n',n'',\cdots} C_n'S_{nn'}\times\sum_{m=2}^{\infty}\Big[\prod_\mu C^0{}_{n^\mu}S_{n^\mu n^\mu+1}\Big]C''{}_{n(m-1)},$$
$$S_{n^\mu n^\mu+1}=\sum_{\sigma^\mu=1}^{N-s^\mu}h(\sigma^\mu)\exp[i\sigma^\mu(1-A)]+\exp[i(n^\mu-n^{\mu+1})\pi]\sum_{\sigma^\mu=1}^{N+s^\mu}h(\sigma^\mu)\times\exp[i\sigma^\mu(1+A)], \quad (6)$$

where

$$C_n'=\epsilon_nA_n\cos n(\theta+\pi/2), \quad C_n''=\epsilon_nA_n\cos n(\alpha-\pi/2),$$

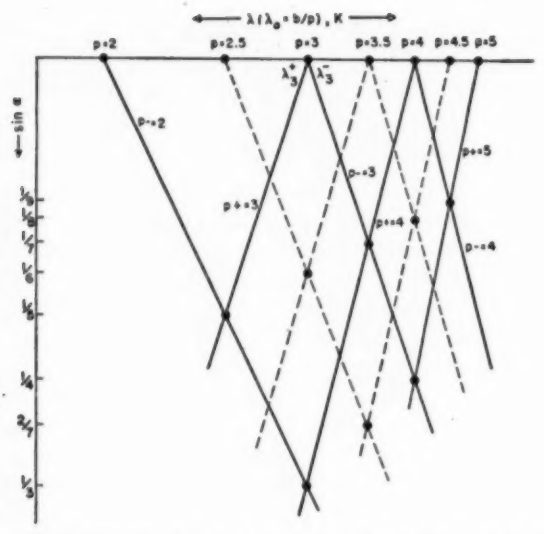


FIG. 2. Schematic of the values of K and $\sin\alpha$ for the individual conditions of (7) and (8) for the first spectral order. The intersections of the heavy lines yield values for which each order of scattering is a maximum, while the intersections of the dotted lines correspond to the minima; points on the individual lines correspond to partial fulfillment of the conditions. The greatest departures from single scattering theory correspond to values obtained by shifting the "network" of heavy lines slightly to the left (each intersection by a distance of the order of b/Np^2), or to slightly longer wavelengths. Compare with Wood, reference 5, Fig. 1.3 (1935), which corresponds to the above with reversed λ axis.

$C_n^{(0)}=\epsilon_n A_n$; for $m=2$ the formal product over μ equals unity.

To this order of approximation in Kb the multiple scattered orders are all symmetrical with respect to the grating. Hence for the analogous grating of semicylindrical bosses on a perfectly conducting plane, departures from single scattering theory will occur only for the component polarized perpendicular to the elements. (See Sec. 2[4] for a discussion of this for an arbitrary planar configuration and also for a discussion of the discrepancies in powers of Kb that (6) entails.)

It can be seen from (6) that for ψ^2, containing $S_{nn'}$, the multiple excitations of the s'th cylinder possess phase factors of the form $\exp[i\sigma(1\pm A)]=\exp[i\sigma Kb(1\pm\sin\alpha)]$; the term in σKb corresponds to the path difference between the s'th cylinder and the sources of its excitation s', while the term in $\pm\sigma Kb\sin\alpha$ is the phase lead or lag of the plane wave excitation at s' relative to that at s. Similarly for ψ^3, possessing $S_{nn'}S_{n'n''}$, the phase factors are of the form $\exp[i\sigma(1\pm A)\pm i\sigma'(1\pm A)]$, where any combination of signs is permissible, etc.

At this point certain interesting conclusions can be drawn in regard to the wavelengths yielding the maximal effects of multiple scattering. (For that matter these follow from rather elementary physical considerations for a grating with arbitrary identical elements subject to the present restrictions on the spacing.) It can be seen from (6) that the phase factors of the S's reduce to unity when

$$Kb(1\pm\sin\alpha)=2p_\pm\pi, \quad p_\pm=1, 2, 3, \cdots, \tag{7}$$

or when $Kb\sin\alpha=2p'\pi$, $Kb=2p\pi$, where p' and p are both integers (p_+ and p_- having even parity), or both half odd integers (p_+ and p_- having odd parity). The physical significance of integral p' is that the primary excitation for these wavelengths is in phase at all cylinders (essentially as for normal incidence), while p integral means that the spacing is an integral multiple of the wavelength; these two conditions insure that the multiple excitations of each cylinder arising from the remaining cylinders are in phase in any particular order of scattering. Similarly for p' and p both half odd integral, although the primary excitation of neighboring cylinders is 180° out of phase, the fact that the spacing is an odd half integral number of wavelengths insures that the excitations of each element in each order are in phase.

Hence if conditions (7) and (5) are fulfilled simultaneously, then the waves scattered by all cylinders in each order of scattering are in phase at P. Thus each order of scattering is a maximum for angles of incidence, wavelengths, and angles of observation specified by $\sin\alpha=(p_+-p_-)/(p_++p_-)$, $\lambda=2b\sin\alpha/(p_+-p_-)=2b/(p_++p_-)$, $\sin\theta=(2q-p_++p_-)/(p_++p_-)$. Now since the magnitude of each order of scattering must be large for any effects of multiple scattering to be appreciable, we expect that the most pronounced, or maximal, effects occur in the vicinities of these values. Whether the intensity is increased or decreased relative to its single scattered value depends on the phase relations among the various orders of scattering. Thus were it possible for all orders to be in phase they would reinforce the first and increase the intensity, while if successive orders could be 180° out of phase they would partially annul the first and the intensity would be less than its single scattered value. In a future section, after we have introduced certain approximations, we will determine the maximal wavelengths for various boundary conditions.

Similarly we deduce conditions for "minimal effects" of multiple scattering, i.e., those for which single scattering theory is most valid. Thus when

$$Kb(1\pm\sin\alpha)=(2p_\pm-1)\pi, \tag{8}$$

the total excitation of each element is a minimum, the initial phase difference at neighboring elements plus the phase difference due to the spacing introducing a total phase difference of 180° between the excitations of each element arriving from successive neighbors. Hence half of the excitations of each element in each order are out of phase with the remaining half. Since this obtains for all orders, regardless of the phases of the orders, we expect the least effects of multiple scattering for the values specified by (8).

In addition to the maximal effects occurring in the vicinity of the wavelengths of (7), we also expect that

fairly large effects will occur in the vicinity of wavelengths satisfying only one of these conditions, i.e., either

$$K^{+}b(1+\sin\alpha)=2p_{+}\pi, \text{ or } K^{-}b(1-\sin\alpha)=2p_{-}\pi, \quad (7')$$

for which some reinforcement among the multiple excitations of each element occurs; e.g., the first insures that waves from successive elements traveling "up" the grating as in Fig. 1 are in phase for ψ^2, etc. These conditions were first proposed by Rayleigh[9] after examining spectrograms obtained by Wood[5] which showed marked departures from the predictions of Fraunhofer theory; Rayleigh later derived these conditions analytically in his dynamical theory of the grating. His interpretation of their physical significance was based on their being the wavelengths having principal maxima in the plane of the grating.

For future convenience Fig. 2 shows the values of K and $\sin\alpha$ for the individual conditions of the sets (7) and (8); the heavy lines correspond to the former, and the dotted lines to the latter, while the circled intersections indicate fulfillment of both conditions of either set. It will be shown later that the actual maximal and minimal wavelengths are displaced only slightly from these positions for the cases to be considered. (It should be noted that Fig. 2 is essentially a schematic of the series of correlated spectrograms obtained by Wood,[5] 1935.)

Although ψ is simplified considerably by substituting the values of (5) and (7) into (6), the interdependence of the limits of the sums over s and σ^{μ} arising from the end effects precludes explicit summation of the ψ^m. To further the analysis we employ the artifice of neglecting the end effects in the following sections.

End Effects Neglected. Explicit Solution

For a grating consisting of an infinite number of elements, the excitation of each cylinder has the same form in any order of scattering regardless of its position in the grating, since each element has just as many elements below it contributing to its excitation as it has above; the excitation of all cylinders are in fact identical for normal incidence. The limits of the sums in (6) would then be $s=-\infty$ to $+\infty$, $\sigma=1$ to ∞, and would be independent of the indices of the preceding sums, there being no end effects for the infinite grating. In the following we assume that the number of elements is very large and take advantage of the simplifying feature of the infinite grating by explicitly neglecting end effects and allowing the sums over σ to range from 1 to N. What we do essentially is to consider the effects at P of only $\mathfrak{N}$ elements of an infinite grating taking into account the contributions to the excitation of each element to all orders of scattering from only its N pairs of nearest neighbors; N is not necessarily $(\mathfrak{N}-1)/2$. By this procedure we simplify the problem considerably but no doubt lose some of the characteristics of the finite grating hinging on the end effects; our motivation is the agreement it yields with experiment.

On neglecting the end effects, (6) reduces to

$$\psi=H_0\Delta(\mathfrak{N})(C+Z), \quad (9)$$

where Z is formally as in (6) but is independent of s. The excitation factor is now

$$S_{n^{\mu}n^{\mu}+1}=\sum_{\sigma=1}^{N}H(\sigma)\{e^{-i\sigma A}+\exp[i(n^{\mu}-n^{\mu+1})\pi]e^{i\sigma A}\}$$
$$\equiv S_{-}+\exp[i(n^{\mu}-n^{\mu+1})\pi]S_{+}\equiv(X, Y), \quad (10)$$

which equals $2\sum H(\sigma)\cos\sigma A\equiv X$ if the indices have even parity and $-i2\sum H(\sigma)\sin\sigma A\equiv Y$ for odd parity. The sum over m as in (6) can now be performed explicitly.

Thus for normal incidence the ψ^m form a geometric progression; the product over μ reduces to $(C_e{}^0X)^{m-2}$, where $C_e{}^0=\sum C_n{}^0$, n even. Hence, provided $|C_e{}^0X|<1$, we obtain

$$\psi=H_0\Delta(\mathfrak{N})[C+C_e'C_e''X/(1-C_e{}^0X)],$$
$$X=2\sum H(\sigma), \quad (11)$$

where C_e', C_e'' are the sums of C_n', C_n'', respectively, for n even and where α equals zero in $\Delta(\mathfrak{N})$, C, and C_e''. Except for X, hitherto untabulated, (11) is no more difficult to employ for calculations than the solution for the single cylinder.‡

It can be seen from the series representation that the ψ^m, $m>1$, are all in phase to reinforce when $\arg(C_e{}^0X)=0$ and that successive ψ^m, $m>1$, are 180° out of phase and partially annul each other when the argument equals π. Whether the single scattered value is effected appreciably, however, depends largely on the magnitude of the sum in X (which has its maximum value of $\sum\sigma^{-1}$ when $Kb=2p\pi$) and also on the phase of the second order relative to the first. Even were the first and second orders in phase, each order would be a maximum and all orders in phase only for the special case where Kb equaled $2p\pi$ and $\arg(C_e{}^0)$ equaled $i\pi/4$ simultaneously. In general, however, because of the particular physical constants of the cylinders, we expect that the greatest effects will be observed when the conditions on the excitations (that they be in phase in each order) and on the orders (that successive orders be in or out of phase) are only partially fulfilled simultaneously.

For $\alpha\neq 0$ the situation is somewhat more complicated; the orders

$$\psi^m=H_0\Delta(\mathfrak{N})\sum_{n,n'\cdots}C_n'S_{nn'}C_{n'}{}^0\cdots$$
$$\times S_{n(m-2)n(m-1)}C''_{n(m-1)}, \quad m>1, \quad (12)$$

‡ We require $|C_e{}^0X|<1$ for the present approximate series representation of (11) to converge. A similar closed form expression for the infinite grating of arbitrary identical elements, however, was obtained subsequently by Dr. S. Karp (through an integral equation approach yielding two linear simultaneous equations for the current density on each element), which is apparently valid for all values of the corresponding quantity.

do not in general reduce to simple products. There are four special cases, however, for which the ψ^m can be summed by inspection, *viz.*, when either of the four quantities Y, X, S_-, or S_+ vanishes. The first case $Y=0$, or $S_-=S_+$, is physically significant since it obtains when the primary excitation is in phase at all elements, fulfilling one of the physical requirements we associate with maximal effects. (The cases $S_+=0$, $S_-=0$ are also of interest since they correspond, respectively, to only excitations traveling either "down" or "up" the grating, as in Fig. 1, being present. These cases, while not particularly realistic for symmetrical elements except for special values of α, may be the more significant ones for gratings with asymmetrical elements or grooves for which pronounced shielding effects may occur in one direction.)

The original derivation of an explicit representation for Z (as in (9)) for arbitrary α was based on the results for these four special cases.§ The procedure was essentially as follows: we construct a general quotient whose terms are polynomials in X, Y which reduces to the known result when either X or Y vanishes; similarly we construct an analogous quotient in S_+, S_-. From the requirement that these two representations, or quotients, be equivalent and by comparison of their corresponding terms, we obtain the closed form for arbitrary α:

$$Z=C_+'Q_-+C_-'Q_+,$$
$$Q_\mp=S_\mp[C_\mp''-(C^0C_\mp''-C^\pi C_\pm'')S_\pm]$$
$$\times\{1-C^0(S_-+S_+)+[(C^0)^2-(C^\pi)^2]S_-S_+\}^{-1}, \quad (13)$$

where

$$C_\pm'=\sum\epsilon_nA_n\cos n(\theta\pm\pi/2), C_\mp''=\sum\epsilon_nA_n\cos n(\alpha\mp\pi/2),$$

$C^0=\sum\epsilon_nA_n$, $C^\pi=\sum(-1)^n\epsilon_nA_n$, and provided that the absolute value of each of the terms of the denominator in $S_\pm$ or S_-S_+ is less than unity. Details will be found in Sec. 4.24, footnote †. A more general derivation by Professor W. Magnus is in the present appendix.

Hence we obtain

$$\psi=H_0\Delta(\mathfrak{N})(C+C_+'Q_-+C_-'Q_+), \quad (14)$$

which constitutes the far field for a grating with width small compared to distance of observation, spacing large compared to wavelength and radius, and subject to the approximation that the end effects have been neglected. The first term of (14) is simply the single scattered wave. The interpretation of the remaining two terms is as follows: the factors $C_\pm'$, or the single cylinder factors with $\alpha=\pm\pi/2$, indicate that these terms consist essentially of two outgoing waves arising from excitations incident on each cylinder in the plane of the grating; the factors $S_\pm$ indicate the distribution of the sources. Each of these terms contain $C_\pm''$, or the single cylinder factor with $\theta=\pm\pi/2$, indicating that these excitations are the waves scattered in the plane of the grating by the remaining cylinders. The presence of C^0 and C^π, or the forward and back scattered single cylinder factors, merely indicates the intermediate effects of intervening cylinders after the plane wave has been scattered by one and before it is finally rescattered. Thus each of the two multiple scattering terms of (14) arises essentially from two exciting waves, one traveling up the grating and the other down.

For an infinite grating the solution would be of the same form as the above but more rigorous in that there would be no need to neglect the end effects. Thus we have

$$\psi=\sum_sH_0(r_s)e^{-is\Delta}\sum_nA_ne^{in\theta_s}(e^{in\alpha}+i^nQ_-+i^{-n}Q_+), \quad (14')$$

where s goes from $-\infty$ to ∞ and σ from 1 and ∞. Although the sum over s can be performed explicitly, we will not consider this case further at present since the finite grating, albeit with end effects neglected, seems a better model on which to base a theory of the Wood anomalies.

III. MODE APPROXIMATIONS

To facilitate further discussion and to derive approximate solutions which can be readily analyzed, we now restrict the problem to radii very small compared to wavelength (or to larger radii provided that the dielectric constant or index of refraction of the cylinders differs little from that of the external medium). The procedure we follow, an extension of that employed to derive (6), retains only the largest terms of all orders of scattering. It is based on the assumption that only certain of the n modes of oscillation of the single cylinder are significant for small values of $x=Ka$. (See Sec. 3[4])

For $x\to0$, the far field solution for the single cylinder to the order of x^2 reduces to $H_0[A_0+2A_1\cos(\theta+\alpha)]$, or the cylinder can oscillate at most in its two lowest modes. Retaining only A_0 and A_1 in all orders, or in Z of (13), yields

$$Z=\{(A_0^2-4A_1^2\sin\theta\sin\alpha)X$$
$$+2A_0A_1[(\sin\alpha-\sin\theta)Y$$
$$+(A_0-2A_1\sin\theta\sin\alpha)(Y^2-X^2)]\}$$
$$\times[1-(A_0+2A_1)X+2A_0A_1(X^2-Y^2)]^{-1}. \quad (15)$$

The "modes" A_0, A_1 and A_0+A_1 will be considered.

For comparison purposes we note that the present second order of scattering equals

$$\psi^2=H_0\Delta(\mathfrak{N})[(A_0^2-4A_1^2\sin\theta\sin\alpha)X$$
$$+2A_0A_1(\sin\alpha-\sin\theta)Y]. \quad (16)$$

Had we not neglected the end effects, we would have obtained a formally identical expression for ψ^2 but with

§ In the original version of this paper only these cases were considered. Equation (13) was obtained after Karp derived his for arbitrary α.

$\Delta(\mathfrak{N})X$, $\Delta(\mathfrak{N})Y$ replaced by

$$2\sum_{s=-N}^{N-1}\sum_{\sigma=1}^{N-s} H(\sigma)\begin{Bmatrix}\cos\\ -i\sin\end{Bmatrix}(s\gamma+\sigma A)$$

$$=2\sum_{\sigma=1}^{2N}\Delta(\mathfrak{N}-\sigma)\begin{Bmatrix}\cos\\ i\sin\end{Bmatrix}(\sigma\gamma'/2),$$

$$\gamma'=Kb(\sin\theta-\sin\alpha), \quad (16')$$

where $\Delta(\mathfrak{N}-\sigma)$ is as in (4') with $\mathfrak{N}$ replaced by $\mathfrak{N}-\sigma$. We will employ this result shortly to compare the magnitudes of ψ^2 for end effects neglected and retained.

A_0 Mode, Isotropic Scatterers

For a plane wave polarized parallel to a cylinder, or for a cylinder of zero impedance in acoustics, and $Ka\to 0$, the largest coefficient is $A_0=iB_0e^{i\phi}$, where B_0 and ϕ are real and positive (as for (10)[4]). Retaining only A_0 in (15) yields

$$\psi=\psi^1/(1-A_0X), \quad \psi^1=H_0A_0\Delta(\mathfrak{N}), \quad X=2\sum H(\sigma)\cos\sigma A. \quad (17)$$

We will first determine the range where single scattering holds and then consider the maximal effects of multiple scattering in detail.[10]

An analytic criterion for the range of single scattering is that the sum of the multiply scattered orders be small compared to the total scattered wave, or that

$$B\equiv|(\psi-\psi^1)/\psi|=|A_0X|=B_02t(2/\pi Kb)^{\frac{1}{2}}\ll 1, \quad (18)$$

where t is defined by

$$te^{i\tau}=T=\sum_{\sigma=1}^{N}\sigma^{-\frac{1}{2}}e^{i\sigma Kb}\cos(\sigma Kb\sin\alpha); \quad (19)$$

T is essentially the sum analog of the Fresnel integral, but its plot in the complex plane is more complicated than the Cornu spiral (see Fig. 4). We note that for the present problem B is also identically $|\psi^2/\psi^1|$. For cases where it is not possible to sum ψ explicitly but where ψ^2 can be readily obtained (e.g., when the end effects are retained), we would also employ $|\psi^2/\psi^1|\ll 1$ as a criterion for single scattering theory; since if $|\psi^2|\ll|\psi^1|$, it follows that $|\psi^3|\ll|\psi^2|$, etc.

It can be seen that the criterion for this case could have been obtained simply by noting that the single scattering solution for the grating, on neglecting the end effects, is A_0X when evaluated at any element. The wave each cylinder scatters in response is then proportional to this excitation factor so that the smallness of the factor insures that the second-order wave scattered by the cylinder is negligible; similarly for the higher orders. This same sort of approach could have been employed to write down (17) immediately; i.e., the excitation of each cylinder in each order is simply a multiple of A_0X so that the wave it scatters is its single scattered wave times this quantity, (17) being simply the sum of all these terms.

The quantity B of (18) differs from that obtained for two cylinders by the factor $2t$. Employing the maximum value of t or

$$\sum_{\sigma=1}^{N}\sigma^{-\frac{1}{2}}$$

for an inclusive criterion yields, e.g., $2t\approx 26$ for $N=50$ indicating a much more limited range of validity than for the two cylinders. Had we employed $|\psi^2/\psi^1|$, with ψ^2 as obtained with (16'), taking the end effects into account, the maximum value of t would be

$$\mathfrak{N}^{-1}\sum_{\sigma=1}^{2N}(\mathfrak{N}-\sigma)\sigma^{-\frac{1}{2}}$$

or $2t\approx 24$ for $N=50$; the difference of the two quantities is negligible for the present purpose. These values of t indicate that the numerical values of Kb and Ka employed to obtain $B<5$ percent for two cylinders now yield that the first- and second-order waves are of the same order of magnitude. Thus for the perfect conductor, $B_0=\pi/2\ln(2/cx)$, $c=1.781\cdots$, we now require $b/\lambda\approx 6\times 10^3$, 2×10^3 for $x=Ka=0.05$, 0.005, respectively, to insure that $B<5$ percent. For two cylinders $b/\lambda\approx 9.4$ sufficed for these values of x. (We ignore a term $\pi^2/4$ compared to $\ln^2(2/cx)$ in B_0 for simplicity.) Similarly for the dielectric, $B_0=\pi x'^2/4$, where for two cylinders $x'=0.3$, $b/\lambda\approx 1$ sufficed, we now require $b/\lambda\approx 600$ for the same value of x', or $x'\approx 0.06$ for the same value of b/λ to insure that $B<5$ percent. (x' equals x times a factor depending on the physical constants of the cylinders; see reference 4.) For larger values of N the effects of multiple scattering become even more pronounced. Thus for $N=5000$ we obtain $2t\approx 280$ (or $2t\approx 260$ on taking into account the end effects) requiring much smaller values of Ka and much larger values of Kb for single scattering theory to be valid.

The minimum values of t, corresponding to the minimal wavelengths of (8), are

$$\sum_{\sigma=1}^{N}(-1)^\sigma\sigma^{-\frac{1}{2}}$$

when the end effects are neglected and

$$\mathfrak{N}^{-1}\sum_{\sigma=1}^{2N}(\mathfrak{N}-\sigma)(-1)^\sigma\sigma^{-\frac{1}{2}}$$

when they are taken into account. For $N=50$ or larger both yield $2t\approx 1$ so that the effects are only of the same order of magnitude as for two cylinders and are, in general, negligible.

In general, the total scattered wave consists of the

[10] A solution which can be put in this form for an infinite grating of isotropically scattering cylinders was obtained previously by W. Franz, Z. ang. Phys. 9, 416 (1949) and R. Honerjäger, Ann. Physik 4, 25 (1948). Their derivation and interpretation were quite different from the present.

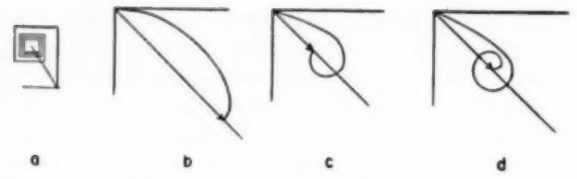


FIG. 3. Sketches of the excitation factor

$$T=\sum_{\sigma=1}^{N}\sigma^{-1/2}e^{i\sigma\delta}=te^{i\tau}=\mathfrak{C}+i\mathfrak{S}$$

in the complex plane for four different values of δ, and N large; the horizontal axis is $\mathfrak{C}$, and the vertical is $\mathfrak{S}$; the variable is N. Case a corresponds to $\tau\approx 125°$, $\delta=\pi/2$, while cases b, c, and d correspond to $\tau=-45°$ and the respective values of $-N\delta\approx 3\pi/4$, $7\pi/4$, and $11\pi/4$.

single scattered wave divided by a rather complicated function of the parameters. The grating scatters isotropically in the sense that a particular wavelength has the same intensity curve in all its spectral orders, or that its principal maxima have the same form. For a particular spectral order, however, although the single scattered intensity varies only slightly with λ and α, the total scattered intensity depends markedly on these quantities. From the discussion of (7) and (8) we expect that a series of collated white light spectrograms for a given spectral order will resemble Fig. 2 in that the most pronounced variations will be found in the vicinities of the heavy lines; however, we have yet to determine their displacement. We now proceed from elementary considerations to deduce the magnitude and direction of the displacement of the "network" of the peak effects from that of Fig. 2, and the qualitative form of the "departure"

$$R\equiv|\psi/\psi^1|^2=|1-A_0X|^{-2}. \tag{20}$$

We rewrite (17) as

$$\psi=\psi^1/(1-Be^{i\beta})=\psi^1[1+Be^{i\beta}+B^2e^{i2\beta}\cdots], \tag{21}$$

where $B=|A_0X|$ as in (18) and where

$$\beta=\tau+\varphi+\pi/4 \tag{22}$$

is the phase difference of successive orders. From elementary considerations we expect intensity maxima when each order of scattering is a maximum and all orders are in phase. The first condition, requiring that excitations in each order are in phase, is met by the wavelengths of (7) which yield $\tau=0$, $t=\sum\sigma^{-\frac{1}{2}}$, while the second requires

$$\beta=2n\pi,\quad n=0,1,2,\cdots,\quad \psi=\psi^1/(1-B). \tag{23}$$

However, these conditions on τ and β can be satisfied simultaneously only when $\phi+\pi/4=2m\pi$, which is a rather unrealistic case. For the usual physical substances no wavelengths satisfy both conditions simultaneously, the maxima occurring when each is only partially fulfilled. (Note that $\phi=0$ for a perfect dielectric and that $\phi=\pi$ for a perfect conductor; the second assumes that $\pi/2\ \ln(2/cx)\ll 1$.)

It is clear that merely requiring that all orders of scattering be in phase, and their graph in the complex plane a straight line, may not necessarily yield maximum scattering since the graph of the excitations may "spiral" and the magnitudes of the individual orders of scattering be small; the extreme case of this is the minimal case (8) mentioned previously. Another example is illustrated in Fig. 3a which is a graph of T in the complex plane for $Kb(1\pm\sin\alpha)=2p_{\pm}\pi+\pi/2$, excitations of each cylinder reaching it from pairs of successive neighboring cylinders (or "successive excitations") being 90° out of phase. It can be seen that although the phase of the resultant τ is roughly 125° so that (23) is approximately satisfied for $\phi=\pi$ (for a perfect conductor), the effects are slight since the magnitude of the resultant t is less than that of the leading term of the series. Similarly for the reverse of this situation, although all excitations in each order may be in phase and their graph a straight line, the graph of the orders of scattering may spiral and the total scattering be small; in fact, if the resultant of the excitations is large while successive orders are out of phase, i.e.,

$$\beta=(2n+1)\pi,\quad \psi=\psi^1/(1-B), \tag{24}$$

minima of the total scattered wave occur.

The graphs of T for the largest values of t for $\tau=-\pi/4$, are sketched in Fig. 3b, c, d for N very large; this value of τ satisfies (23) for $\phi=0$, the perfect dielectric, and (24) for $\phi=\pi$, the perfect conductor. The largest of these values of t occurs when the phase difference of successive excitations, say δ, is sufficiently small so that the graph approximates a "simple arc" as in Fig. 3b. This value of δ, say δ', yields the extrema of $(1\mp B)^{-1}$ corresponding to an intensity decrease for the conductor, and an intensity increase for the dielectric. We expect, however, that the maximal effects will occur for a smaller value of $|\delta|$, say $\delta_\wedge$ lying between 0 and δ', for which the excitations are more nearly in phase than is required by (23), (24). Thus $\delta=0$ satisfies (7), while $\delta=\delta'$ satisfies (23), (24); neither by itself yields the maximal effects, which occur rather for, say, $\delta_\wedge\approx\delta'/2$, as a rough figure of merit.

We obtain an approximate upper limit or figure of merit for δ', in terms of N and ϕ, from the following considerations.|| The largest value of t with a phase of $-45°$ occurs when the graph of the excitations is still a simple arc as in Fig. 3b. For this simple arc case to obtain, we require that the last vector (the N'th excitation) make an angle less than 90° with the resultant, since if this angle is greater than 90° the graph has already begun to "spiral" to assume the form of Fig. 3c and the resultant has begun to decrease. Hence, for the largest value of t commensurate with $\tau=-45°$ we have $|N\delta'|\leq\pi/4+\pi/2$. In general, $|N\delta'|\leq|\tau|+\pi/2$ where $|\tau|$ is prescribed by (23), (24).

|| The writer is indebted to Professor W. Magnus for suggesting this.

We accept $|\delta'| \approx 3\pi/4N$ as a convenient figure of merit. The most interesting feature of this result is that δ' depends only on N, the number of nearest neighbors we take into account, and on ϕ, the phase of the scattering coefficient of the single cylinder, and is apparently independent of the remaining parameters of the problem. Now since the maximal value δ_Λ does not differ greatly from δ', we accept as a working hypothesis that δ_Λ also depends primarily on N and ϕ and is independent of the other parameters. Hence the greatest effects, decreasing the intensity for the perfect conductor and increasing it for the perfect dielectric, occur for the maximal values

$$K_\Lambda b(1 \pm \sin\alpha) = 2p_\pm \pi + \delta_\Lambda, \quad T_\Lambda = \sum \sigma^{-\frac{1}{2}} e^{i\sigma\delta_\Lambda}, \quad (25)$$

where δ_Λ is a small negative quantity between 0 and $-3\pi/4N$. From this expression certain interesting features of the intensity curves can be obtained immediately.

On writing $K_\Lambda b = K_0 b + \delta_\Lambda$, where K_0 corresponds to the wavelengths for $\delta = 0$, we obtain for the displacement of the maximal wavelengths from the values of (7)

$$\Delta_\Lambda = \lambda_\Lambda - \lambda_0 = -\delta_\Lambda \lambda_0 \lambda_\Lambda / 2\pi b \approx |\delta_\Lambda| b/2\pi p^2 < 3\lambda_0^2/8Nb. \quad (26)$$

This indicates that the maximal effects are "shifted to the red," i.e., occur on the long wavelength side of the wavelengths $\lambda_0 = b/p$, the displacement being roughly proportional to λ_0^2/Nb, the square of the wavelength divided by the width of the grating. (For the conducting dielectric, $\Delta_\Lambda < (\phi + 3\pi/4)/2\pi bN$.) From this and the preceding discussion it can be seen that the most pronounced effects cover a finite portion of the λ axis, and form a "maximal band" whose "peak" is at $\lambda_\Lambda > \lambda_0$; this agrees with references 5–8. The band widths are essentially proportional to Δ_Λ so that their ratio for different λ_Λ's, the other parameters being fixed, is approximately

$$w_p/w_{p+1} = w \approx (p+1)^2/p^2. \quad (27)$$

Thus $w \approx 1.8$ for $p = 3$ or the band near $\lambda_0 = b/3$ should be almost twice as wide as that near $\lambda_0 = b/4$; this agrees with spectrograms for $\alpha = 0$ obtained by Wood,[5] 1935.

If the incident radiation is white, we expect that these bands will be dark relative to neighboring wavelengths if the elements are perfectly conducting, and bright if the elements are perfect dielectrics. It is evident from Fig. 3, however, that these bands are not simple intensity extrema (as for two cylinders), but that they are flanked, at least on the long wavelength side, by a succession of "local extrema" which become less pronounced as $|\delta|$ increases. From Fig. 3 we can deduce the form of the curve of $R = |\psi/\psi^1|^2$ *vs* λ, or R *vs* $-\delta$. Thus for the conductor as λ increases, the curve passes through a deep minimum indicated by Fig. 3b, rises to a low local maximum with R still less than unity indicated by Fig. 3c, and falls to another minimum not as low as the first as indicated by Fig. 3d. With increasing λ the values of these local extrema approach unity, or $\psi \rightarrow \psi^1$, as λ approaches its minimal value, or $\delta \rightarrow \pi$. The approximate positions of the first few local extrema are obtained by adding multiples of π/N to that for the maximal value; i.e., $N\delta' \approx -\pi/4 - (2n-1)\pi/2$, the odd values of n corresponding to the minima, and the even to the maxima. Only the first few are significant, however, because of the behavior of T with increasing δ; i.e., it can be deduced that T has the spiral form indicated by Fig. 3d until $|\delta| = \pi/2$ is reached, and that thereafter the graph becomes a jagged "star patch" with a rapidly decreasing resultant. The largest value of $|\delta|$ yielding $\rho = -45°$, however, is much less than $\pi/2$, and the largest value of $|\delta|$ yielding a resultant of appreciable magnitude is smaller still.

For the perfect dielectric the situation is essentially as above and its discussion requires merely the interchange of the words maximum and minimum, high and low, etc.

Another significant feature of the $R-\lambda$ curve that can be established without numerical calculation is that the sharp boundary of the peak occurs on its short wavelength side. Thus for $A = 2p'\pi$ the largest term of $\partial R/\partial\lambda$ is $(bK^2R^2B/\pi l)\sum(\sqrt{\sigma}) \sin(Kb + \phi + \pi/4)$. The largest value of this term occurs in the immediate vicinity of λ_0 where this quantity, $\propto g = \sin(\phi + \pi/4)\sum(\sqrt{\sigma})$, is very large. For the perfect conductor $\phi = \pi$, g equals $-|g|$ indicating that the intensity drops sharply in the vicinity of λ_0, while for the perfect dielectric $\phi = 0$, g equals $|g|$ indicating a sharp intensity rise. (The sharp boundary at λ_0 agrees with the previously mentioned experiments.)

For $\lambda < \lambda_0$, $\delta = |\delta|$ (the short wavelength side of the peak), the situation is not as clear cut. Although the value $\tau = 3\pi/4$ satisfies (23), (24), yielding opposite effects to those for $\tau = -\pi/4$, these are negligible since T falls in the "star patch" region mentioned previously. We surmise, however, that other positive values of τ exist for which the effects are appreciable, since the $R-\lambda$ curve for $\lambda < \lambda_0$ should be essentially of the same form as for $\lambda > \lambda_0$. The two regions of the curve must be complimentary in character, however, in order for energy to be conserved. Furthermore, since the "periodicity" of the extrema is $2\pi/N$ in the vicinity of λ_0, the local extrema for $\lambda < \lambda_0$ correspond to $N\delta' \leq (1/4 + n)\pi$, so that they interlace those for $\lambda > \lambda_0$, the first being shifted more to the blue than the peak was to the red. These values of δ', corresponding to the last vector roughly parallel to the 45° line, yield a set of positive values of τ between $\pi/8$ and $3\pi/8$ which approach $\pi/4$ with increasing n. It is not obvious, however, that these values of τ yield local extrema for $\lambda < \lambda_0$ since successive orders of scattering are neither approximately in nor out of phase. While the magnitudes and phases of the orders may be such as to yield appreciable departures of R from unity, this remains to be established numerically.

We return now to consideration of the maximal peak, to obtain some idea of its angular width. We take as a figure of merit the range bounded by λ_0 and the first local extremum on its long wavelength side, i.e., $(7\pi/4N)\times(b/2\pi p^2)\approx b/Np^2=W$. For normal incidence the angular positions of the λ_0's are given by $\sin\theta=q/p=q\lambda_0/b$. Hence $\Delta(\sin\theta)=q\Delta\lambda_0/b$, or for relatively small values of θ, $\Delta\theta\approx qW/b=q/Np^2$. Thus the maximal band is quite narrow, subtending only a small fraction of a degree, even for only moderately large values of N; e.g., for $q=1$, $p=3$, and $N=10$, 10^2, 10^3, 10^4 we obtain $\Delta\theta\approx 40'$, $4'$, $0.4'$, $0.04'$, respectively. In general, we have $\Delta\theta=q/Np^2\cos\theta=q\lambda_0/bpN\cos\theta$ which for $\mathfrak{N}\approx 2N$ is of the same order of magnitude as $\lambda_0/\mathfrak{N}b\cos\theta$, the angular half-width of a principal maximum of $\Delta(\mathfrak{N})$. It can be seen that if all parameters but α and θ are held constant (the bands occurring for $\sin\alpha=p'/p$, $\sin\theta_0=(q-p')/p$, $p'=0, 1, 2\cdots p$), then as α increases, θ decreases and $\Delta\theta$ or the band width should decrease. (See Wood,[5] 1935, where this decrease occurs but more rapidly than if due solely to a $\sec\theta$ effect.) The resolving power for the maximal band is $\lambda_\wedge/\Delta_\wedge\approx pN$, or of the same order of magnitude as $q\mathfrak{N}$ for a principal maximum of $\Delta(\mathfrak{N})$. This implies that for a discrete source, multiple scattering may have marked effects on one of two closely neighboring lines without effecting the other appreciably. (This was observed by Wood,[5] 1902, with the sodium D lines having a separation of 6A or $\lambda/\Delta\approx 10^3$.)

It should be kept in mind that although N is proportional to $\mathfrak{N}$, the number of elements, it is not necessarily of the same order of magnitude, particularly for gratings with more complicated elements for which shielding effects arise, or when absorption in the external medium is significant. These present relations, however, when correlated with measurements of the widths of the peaks and their displacement from λ_0 should indicate an appropriate value of N for a particular grating.

The relative intensities of the bands in different parts of the spectrum can also be deduced. For the perfect conductor, $B=t/p^{\frac{1}{2}}\ln(2/cx)\propto\lambda^{\frac{1}{2}}/\ln(\lambda/ca\pi)$ since we assume that $\delta_\wedge$, and consequently $t_\wedge$, is independent of the wavelength; hence B increases as λ increases, the effect of the square root overshadowing that of the logarithm. The bands should therefore depart more from their single scattered values for the longer wavelengths. These are not necessarily the darkest, however, since the total intensity increases as $\lambda/\ln^2(\lambda/ca\pi)$. The least intense bands (lowest values of $|\psi|^2$) may therefore occur for the shorter wavelengths, while the most pronounced (lowest values of R) should occur for the longer wavelengths.

For the perfect dielectric, $B=x'^2t/2p^{\frac{1}{2}}\propto\lambda^{-\frac{1}{2}}$; thus B increases rapidly with decreasing wavelength. Hence the largest values of R occur for the shorter wavelengths and these bands are also the brightest since $|\psi|^2\propto\lambda^{-3}$. For the conducting dielectric the situation is essentially as for the perfect dielectric, but since for this case ϕ is usually between 0 and $\pi/2$ we expect that $\delta_\wedge$ will be somewhat larger [corresponding to the larger values of τ prescribed by (23)]. Consequently the maximal effects occur for longer wavelengths and are less pronounced than for the perfect dielectric.

It can also be seen that although the bright bands may have intensities many times their single scattered values, the least intensity of a dark band is only of the order of one-fourth its single scattered value (since all orders reinforce the first for the maximum, while there is only a partial annulment for the minimum). This is of course evident from ψ of (24) and the condition that $B_{\max}{}^2<1$ to insure convergence.

We have for simplicity confined ourselves to this point to the bands for which $A=Kb\sin\alpha=2p'\pi$, occurring in the vicinity of and displaced to the lower left of the intersections of the heavy lines in Fig. 2. (Those of the figure occur at $\lambda_0=b/p$, $\sin\alpha=p'/p$, whereas the ones we considered fall at longer wavelengths, $\lambda=\lambda_0+\Delta_\wedge$, and at very slightly larger angles, $\sin\alpha=2p'\pi/(2p\pi+\delta_\wedge)$.) The discussion will now be extended to the partial maximal effects corresponding to the heavy lines of Fig. 2 and the conditions of (7′).

We write

$$2T=T_++T_-=t_+e^{i\tau_+}+t_-e^{i\tau_-}=\sum\sigma^{-\frac{1}{2}}\{e^{i\sigma(1+A)}+e^{i\sigma(1-A)}\},$$
$$B_\pm=B_0(2/Kb\pi)^{\frac{1}{2}}l_\pm,\quad \beta_\pm=\varphi+\tau_\pm+\pi/4,$$

and obtain in analogy to (21) and (23), (24):

$$\psi=\psi^1/(1-B_+e^{i\beta_+}-B_-e^{i\beta_-}), \tag{28}$$

and

$$\beta_\pm=2n_\pm\pi,\quad =(2n_\pm+1)\pi. \tag{29}$$

Thus if $\beta_+=2n_+\pi$ while β_- is arbitrary, the scattering is increased, while if $\beta_+=(2n_++1)\pi$, it is decreased. For wavelengths in the vicinity of K^+ we write

$$\begin{aligned}Kb(1+\sin\alpha)&=2p_+\pi+\delta^+=K^+b(1+\sin\alpha)+\delta^+,\\ Kb(1-\sin\alpha)&=(2p_+\pi+\delta^+)(1-\sin\alpha)/(1+\sin\alpha),\end{aligned} \tag{30}$$

where we have omitted the maximal symbol $\wedge$ for brevity, and obtain

$$2T=\sum\sigma^{-\frac{1}{2}}\{\exp(i\sigma\delta^+)+\exp[i\sigma(2p_+\pi+\delta^+)(1-\sin\alpha)/(1+\sin\alpha)]\}, \tag{31}$$

which reduces to $2T$ as in (25) for $Kb\sin\alpha=2p'\pi$, $\delta^-=\delta^+=\delta$. (Similarly for the vicinity of K^-, we replace $\pm$ by $\mp$ in the above.) It can be seen that in general the effects now depend on λ and α. Thus if $Kb(1-\sin\alpha)$ differs appreciably from $2p\pi$, the contribution of T_- are small and B is of the order of half the value it has at the intersections; while if $Kb(1-\sin\alpha)\approx 2p\pi$, the effects are essentially as for the intersections.

For the shift of the peaks to longer wavelengths corresponding to (26) we have

$$\Delta_\pm=\lambda_\wedge{}^\pm-\lambda^\pm=-\delta^\pm\lambda_\wedge{}^\pm\lambda^\pm/2\pi b(1\pm\sin\alpha). \tag{32}$$

We note that for the case treated previously, $A=2p'\pi$, although $\delta^\pm=\delta_\wedge$, the peak values corresponding to λ^+ and λ^- were displaced by different amounts in order for

them to coincide at $\lambda=\lambda_0+\Delta_\wedge$; i.e., $\Delta_+\neq\Delta_-\neq\Delta_\wedge$ so that for a given value of α the intersection at $A=2p'\pi$ arises for $\lambda^+\neq\lambda^-$. The peaks corresponding to the intersections of the heavy lines of Fig. 2 ($\lambda^+=\lambda^-=\lambda_0$, $\sin\alpha=(p_+-p_-)/(p_++p_-)$, $A\neq 2p'\pi$) are shifted by $\Delta_+=\Delta_-\equiv\Delta_0$, but $\delta^+\neq\delta^-$ and there is no displacement in α; thus if $Kb(1\pm\sin\alpha)=2p_\pm\pi+\delta^\pm=K^\pm b(1\pm\sin\alpha)+\delta^\pm$, $K^\pm=K_0$, we have $\Delta_\pm=\Delta_0=-\delta_0\lambda_\wedge\lambda_0/2\pi b$, $\delta_0=\delta^\pm/(1\pm\sin\alpha)$. These quantities Δ_0, δ_0 should not be confused with those occurring for $A=2p'\pi$, although the values of λ and α at which the two effects occur are practically identical. We note that for the present case $2Kb=2K_0b+2\delta_0$, formally analogous to the previous, but that $2Kb\sin\alpha=4p'\pi+2\delta_0\sin\alpha$. This differs from the previous by $2\delta_0\sin\alpha$ which indicates that the effects are at the same value of α at which λ_0 occurs rather than at a larger value of α as for the previous case.

Introducing $\delta^\pm=\delta_0(1\pm\sin\alpha)$ into (32) and assuming as for the previous $\delta_\wedge$ that δ_0 is independent of the wavelength yields for the partial maximal effects

$$\Delta_\pm=-\delta_0\lambda_\wedge{}^\pm\lambda^\pm/2\pi b\approx-\delta_0(1\pm\sin\alpha)^2b/2\pi p_\pm{}^2, \quad (32')$$

so that the shift to the red of the peaks of the bands and the band widths are proportional to the square of the wavelength as for the case treated previously.

Substituting $\delta^\pm=\delta_0(1\pm\sin\alpha)$ into T of (31) and into its analog for K^- yields

$$2T=\sum\sigma^{-\frac{1}{2}}\{\exp[i\sigma\delta_0(1\pm\sin\alpha)]+\exp[i\sigma\delta_0(1\mp\sin\alpha)+i\sigma 2p_\pm(1\mp\sin\alpha)/(1\pm\sin\alpha)]\}. \quad (31')$$

This reduces to

$$2T=\sum\sigma^{-\frac{1}{2}}\{\exp[i\sigma\delta_0(1+\sin\alpha)]+\exp[i\sigma\delta_0(1-\sin\alpha)]\}$$

for $\lambda^+=\lambda^-$, where $(1\mp\sin\alpha)/(1\pm\sin\alpha)=p^\pm/p_\pm$, which although not as simple as (25) (except for $\alpha=0$) is of the same order of magnitude and indicates essentially the same sort of effects.

Numerical analysis: The results of numerical investigation of the total scattered wave of (21) will now be presented. This work was undertaken to support the above physical arguments and to check the assumption that $\delta_\wedge$ depended primarily on N and ϕ and only slightly on λ. The specific (and not necessarily realistic) numerical examples considered show that these arguments and this assumption were justified, and also substantiate the various figures of merit and relations for the ratio of the band widths, their structure, resolutions, etc., that were obtained.

Employing the formalism of (25) for all values of λ for $A=2p'\pi$ we write

$$T=\sum\sigma^{-\frac{1}{2}}e^{i\sigma\delta}\equiv\mathfrak{C}+i\mathfrak{S}, \qquad \delta=Kb(1\pm\sin\alpha)-2p_\pm\pi=Kb-2p\pi, \quad (33)$$

where $\mathfrak{C}$ and $\mathfrak{S}$ are real. Employing the "Euler sum formula"[11] to derive a more rapidly convergent representation for $\mathfrak{C}$ and $\mathfrak{S}$ yields

$$\mathfrak{C}=(2\pi/\delta)^{\frac{1}{2}}[C(N\delta)-C(\delta)]+(\cos\delta+\cos N\delta/\sqrt{N})/2+(\cos\delta/2+\delta\sin\delta-\cos N\delta/2N^{\frac{1}{2}}-\delta\sin N\delta/\sqrt{N})/12\cdots,$$
$$\mathfrak{S}=(2\pi/\delta)^{\frac{1}{2}}[S(N\delta)-S(\delta)]+(\sin\delta+\sin N\delta/\sqrt{N})/2+(\sin\delta/2-\delta\cos\delta-\sin N\delta/2N^{\frac{1}{2}}+\delta\cos N\delta/\sqrt{N})/12\cdots, \quad (34)$$

where $C(N\delta)$, $S(N\delta)$, etc., are the tabulated Fresnel integrals.[12] These are rapidly convergent for N large or small. Thus for $\delta=0$, $\sum\sigma^{-\frac{1}{2}}=\mathfrak{C}\approx 2N^{\frac{1}{2}}-1.46+\frac{1}{2}N^{-\frac{1}{2}}$ yields results agreeing to at least three placed for $N>50$ with those obtained by adding tabulated values of $\sigma^{-\frac{1}{2}}$; e.g., 18.6, 61.7, 198.5, for $N=10^2$, 10^3, 10^4, respectively. For $\delta=\pi$, (34) yields an error of only about 5 percent for large N.

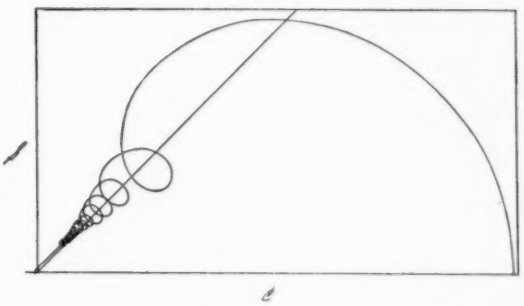


FIG. 4. Sketch of T in the complex plane for $N=10^4$ and δ varying from 0 to π.

A plot of $\mathfrak{C}$ *vs* $\mathfrak{S}$ for $\delta=0$ to π resembles the sketch in Fig. 4 for large N; the "pigtail" spiral becomes more or less parallel to the 45° line with increasing N, but the curves for different values of N cannot be made to coincide by simply changing the scale factors.¶ (For $\delta=0$ to $-\pi$ we obtain the image of this curve in the real axis.) The periodicity of this spiral does not differ greatly from that of the Cornu spiral, as expected, and the present curve is what would be "obtained" were the starting point of the Cornu spiral in the first quadrant shifted to $\sum\sigma^{-\frac{1}{2}}$ along the real axis and its center pulled out along the 45° line to about -0.5. It can be seen that Fig. 3b, c, and d correspond, respectively, to the first three intersections of the image of this curve with the $-45°$ line, while the local extrema for $\lambda<\lambda_0$ correspond to portions of the curve parallel to the $+45°$ line. The first intersection occurs for $|\delta|\approx 2.9\pi/4N$ for $N=10^2$, 10^3, and 10^4, which agrees with our previous upper limit $|\delta'|\leq 3\pi/4N$.

Figures 5, 6, and 7 are of $R=|\psi/\psi^1|^2$ *vs* $\delta=Kb-2p\pi$ in degrees for $N=10^4$, 10^3, and 10^2, respectively. The *a* curves, the dark bands, are for the perfect conductor

[11] See, for example, L. Bieberbach, *Lehrbuch der Funktionentheorie* (Chelsea Publishing Company, New York, 1945), Vol. I, p. 307.

[12] E. Jahnke and F. Emde, *Tables of Functions* (Dover Publications, New York, 1945), p. 35.

¶ This function has been analyzed graphically and numerically with N and δ as parameters; tables and curves will appear shortly in a report by Joy Russek and the writer. Mrs. Russek was also of great assistance in the numerical computations required for the graphs of the present paper.

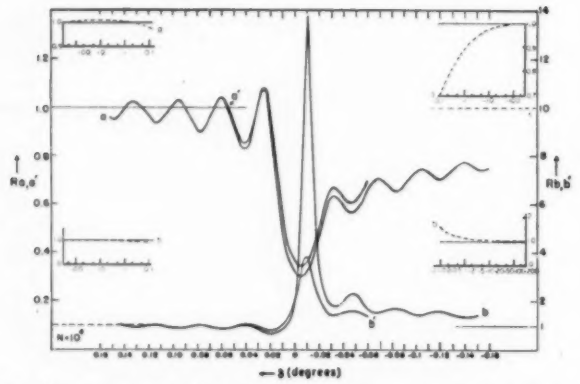


FIG. 5. Plots of $R=|\psi/\psi^1|^2$ *vs* $\delta=Kb-2p\pi$ for the polarization component parallel to the elements of the transmission grating with Ka very small; computations for the curves are accurate to about 5 percent. The dark bands a, a' are for the perfect conductor for $\lambda_0=b/3$, $b/4$; the left R scale is used for these curves. The bright bands b,b' are for the perfect dielectric for $\lambda_0=b/4$, $b/3$; the right R scale is used. The displacement of the curves from the horizontal lines $R=1$ indicates the departure from single scattering theory. The inserts show mean values of R on logarithmic δ scales. Because of the approximations employed (particularly the neglect of the end effects), only the gross features of these curves should be observed experimentally. The present curves are for $N=10^4$; values of the other constants are in the text.

for $p=3$, and the a' curves or their remnants are for $p=4$; the left R scale is to be used for these curves. The b curves, the bright bands, are for the perfect dielectric for $p=4$, and the b' curves or their remnants are for $p=3$; the right R scale, with a scale factor one-tenth that of the left, is to be used for these curves. Note the reversal of the wavelengths $\lambda_0=b/3$, $b/4$ for the primed and unprimed curves for $\phi=0$ and π. The unprimed curves, corresponding to the wavelengths yielding the larger effects for the particular case, are associated with the longer wavelength for the conductor and with the shorter for the dielectric.

To facilitate comparison of the bands for different values of N, the same "normalization," $B_{\max}=B_0 2\sum\sigma^{-1}/\pi p^1=0.9$ was chosen for the unprimed curves for all N so that all have the same value of R for $\delta=0$. This prescribes different values for the ratio of the radius to spacing, say u, for each value of N and B_0, u increasing markedly with decreasing N, particularly for the perfect conductor. Thus for curves a and a' we have $u=2.53\times10^{-57}$, 3.77×10^{-19}, 3.93×10^{-7}, while for curves b and b', $u'=5.37\times10^{-3}$, 9.62×10^{-3}, 1.75×10^{-2}, for $N=10^4$, 10^3, 10^2, respectively. Hence the larger the value of u, the smaller the value of N required to obtain intensity ratios of the same order of magnitude. The normalization also prescribes the maximum value of B for the primed curves so that for the a' curves $B_{\max}=0.782$, 0.784, 0.798, for $N=10^4$, 10^3, 10^2, while for the b' curves $B_{\max}=0.585$ for all N.

Our surmise that the maximal value of δ depended primarily on N for a given value of ϕ is substantiated by these curves; there is only a very slight shift to a smaller value of $\delta_\wedge$ for the larger λ_0. Although corresponding curves of the graphs are not congruent, their deviations are relatively superficial and the positions of the local extrema agree closely on a scale of δ/N. Thus the peaks of the maximal bright bands occur for $\delta_\wedge\approx-(2.4)\pi/4N$, while the "peaks" of the dark bands fall at $\delta_\wedge\approx-(1.1)\pi/4N$. It can also be seen that the largest absolute slopes occur for the short wavelength side of the peaks as deduced previously.

Our remarks on the character and positions of the local extrema are valid in their essentials. Although the departures of R for $\lambda>\lambda_0$ are more pronounced than those of their complements for $\lambda<\lambda_0$, the former approach unity more rapidly as $\delta\to\pi$. Taking into consideration the logarithmic scales of the inserts (showing mean values of R for $|\delta|$ approaching the minimal value of 180°) we surmise that a more detailed investigation will show that energy is conserved in the process. The distance between the local extrema flanking the peak is approximately 550°/N, or differs by less than a factor of 2 from the value $8\pi/4N$ employed previously as a figure of merit. The peak of the bright band, however, is narrower than that of the dark and is also more pronounced because of its relatively less prominent local extrema. The "contrast" of the bands, or $R_{\max}/R_{\min}$, increases with increasing λ_0 for the conductor and decreases for the dielectric, the change being much more pronounced for the latter. The contrast also increases, and the peaks become more marked, with decreasing N for the same normalization.

These curves for R correspond physically to values obtained for fixed α and θ with incident monochromatic radiation of different wavelengths, the suppressed graph of $|\psi^1|^2$ being that of the principal maximum and secondary maxima of $\Delta(\mathfrak{N})$ for these angles. However, we may also regard these graphs, and those of $|\psi|^2$ to be discussed shortly, as corresponding to values obtained for fixed α but with varying monochromatic radiation and with varying θ, the angle θ being that which yields a principal maximum for the particular λ and α under consideration. Hence the single scattered intensity, equaling $|\psi^1|^2=\mathfrak{N}^2B_0^2 2/\pi Kr$, can be considered as constant in the immediate vicinity of λ_0 and identical for curves a and b by virtue of the normalization.

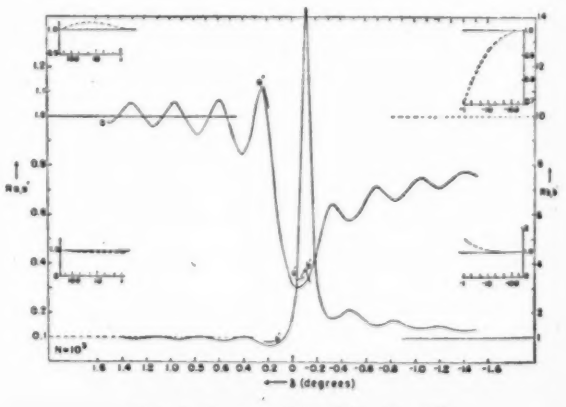


FIG. 6. Same as Fig. 5 for $N=10^3$.

Were we to ignore possible discrepancies arising from the neglected end effects, the mode hypothesis, and the presence of the transmitted plane wave, and were we to assume in addition that the spectral lines of $\Delta(\mathfrak{N})$ were of infinitesimal width, we could regard these curves as corresponding to spectrograms in the vicinity of the λ_0's for incident white light. Because of the finite width of a principal maximum of $\Delta(\mathfrak{N})$ and because of its secondary maxima, however, there will always be present some "background" light of different wavelength than that corresponding to a particular angle of observation. To facilitate further discussion and comparison with experiment, however, we accept these curves as corresponding to spectrograms for white light, but expect that many of their details will not be detected experimentally.

Figure 8, showing $|\psi|^2$ *vs* λ for the parameters of Fig. 6 (unity on the intensity scale equaling $|\psi^1|^2$ for $p=3$, $\phi=\pi$, or $p=4$, $\phi=0$), is therefore regarded as approximating an experimental intensity curve near the λ_0's. We expect, however, that rather than our oscillatory bands, "wedges" bounded sharply on their short wavelength side and "tailing off" on the long wavelength side will be observed experimentally. The widths of the peaks are $1.025\times10^{-5}b$ for $p=4$ and $1.82\times10^{5}b$ for $p=3$, their ratio being 9/16 as determined previously; the bands with their tails are of course much wider, but the ratio should be preserved. The darker band occurs for the shorter wavelength for the conductor and for the longer wavelength for the dielectric.

The quantity $|\psi|^2$ can be considered as proportional to the total intensity or Poynting flux provided that $\Delta(\mathfrak{N})$ is sufficiently large so that the incident wave and the interference terms of the incident and scattered waves can be neglected. For the minima of $|\psi|^2$, however, or in general when $\mathfrak{N}$ is small, these must be taken into account as in (21),[4] (22)[4] to obtain the total physical fields and derived quantities for comparison with experiment. To evaluate these quantities we require

$$f=Re[\psi e^{-iKr\cos(\theta+\alpha)}]=DR(\cos q-B\cos q'), \quad (35)$$

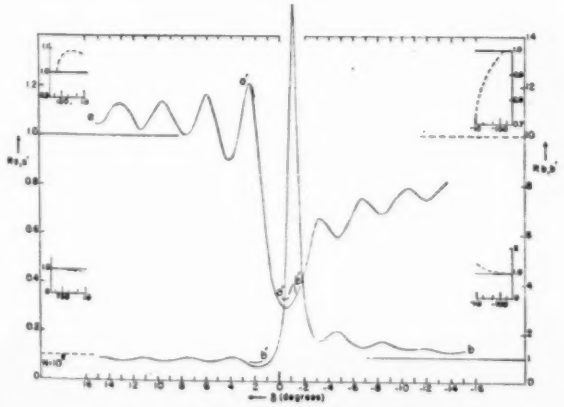


FIG. 7. Same as Fig. 5 for $N=10^6$.

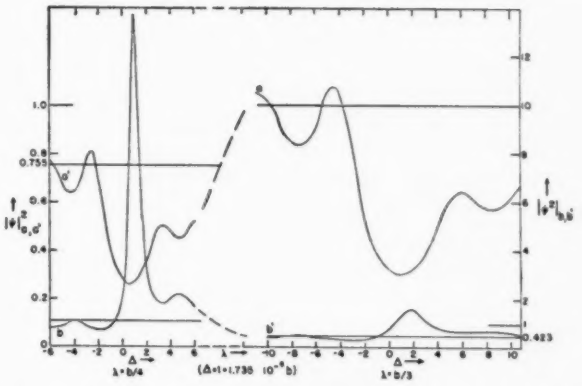


FIG. 8. Plots of $|\psi|^2$ *vs* λ for the values of Fig. 5. The left $|\psi|^2$ scale is for curves a,a'; the right, for b,b'. The normalization is $|\psi^1|_a=|\psi^1|_b=1$.

where

$$D=\Delta(\mathfrak{N})B_0(2/\pi Kr)^{\frac{1}{2}}, \quad R=|1-Be^{i\beta}|^{-2}$$
$$q=Kr[1-\cos(\theta+\alpha)]+\phi+\pi/4,$$

and

$$q'=q-\beta=Kr[1-\cos(\theta+\alpha)]-\tau.$$

For $D\gg0.1$ these terms are negligible compared to $|\psi|^2=D^2R$, particularly for R large. For moderate values of D and small values of R, however, they must be considered. Thus when successive orders are out of phase, $f=-D\cos\{Kr[1-\cos(\theta+\alpha)]+\pi/4\}/(1+B)$ and $|\psi|^2=D^2/(1+B)^2$. The largest values of B are of the order of unity so that the first term is proportional to $D/2$ as compared to $D^2/4$ for the second. If these two are of the same order of magnitude, or if the first is much greater than the second, the intensity curves will be quite different from those we have considered. More generally, the plane wave and interference terms will tend to smooth out the local extrema to form a gradual tail for the effects.

We note that the convergence requirement, $B_{\max}{}^2<1$, for a given value of N, λ, and b/r furnish an upper limit for the single scattered intensity, or D^2, provided that N is the total number of pairs of cylinders (rather than some fraction). Thus since $|\psi^1|^2=D^2=B_{\max}{}^2(N/t_m)^2\times(b/r)<(N/t_m)^2(b/r)$ and since we require that $Nb\ll r$, we have $D^2\ll N/t_m{}^2$. (We have eliminated B_0 by means of (18), replaced $\Delta(\mathfrak{N})$ by $2N$, and written t_m for the maximum value of t.) Using the previous values for $t_m=\sum^{\sigma-\frac{1}{2}}$ we have $t_m{}^2\approx4N$. Hence $D^2\ll1/4$ so that the plane wave and the interference term need to be taken into account, particularly for the minima. It can be seen that the convergence requirement necessitates that the usual Fraunhofer effects be weak for these gratings if N is of the order of $\mathfrak{N}/2$.

A_1 Mode

For a plane wave polarized perpendicular to the axis of a dielectric cylinder, or a gaseous cylinder in acoustics, and $Ka\to0$, the largest coefficient is $A_1=iB_1e^{i\phi}$ (as for (16)[4]). Retaining only this coefficient in (15)

yields

$$\psi=\Delta(\mathfrak{N})H_0 2A_1[\cos\theta\cos\alpha-\sin\theta\sin\alpha/(1-2A_1X)]$$
$$=\psi_e{}^1+\psi_0{}^1/(1-2A_1X)=\psi_e+\psi_0, \quad (36)$$

where $\psi^1=\psi_e{}^1+\psi_0{}^1=H_0\Delta(\mathfrak{N})2A_1\cos(\theta+\alpha)$. To this order of approximation only the odd component, or that symmetrical to the grating, is effected by multiple scattering, the solution reducing to its single scattered value for either α or $\theta=0$. With increasing α and θ the odd component prevails (plots of R for $\psi_e\approx\psi_0$ would differ greatly from those discussed previously), until finally for either α or $\theta=\pi/2$ the total scattered wave becomes formally identical with ψ of (17) for the dielectric case.

Hence for large α and θ for a perfect dielectric, maximal effects of the same type occur for both the perpendicular and parallel components and at the same wavelengths. However, since the present $2B$ equals $2/(k_e+1)$ times the B that occurred for the parallel component,** and also because of the "background" factor $\cos\theta\cos\alpha$, the effects are now less pronounced. For the conducting dielectric the phase factors ϕ are no longer zero and maximal effects occur at different wavelengths for the two components.

The total scattered intensity and the interference terms are given by

$$|\psi|^2=4D^2\{\cos^2\alpha\cos^2\theta-R_0\sin\theta\sin\alpha$$
$$\times[2(1-2B\cos\beta)\cos\theta\cos\alpha-\sin\theta\sin\alpha]\}, \quad (37)$$
$$f=2D\{\cos\theta\cos\alpha\cos q-R_0\sin\theta\sin\alpha(\cos q-2B\cos q')\}$$
$$R_0=|1-2Be^{i\beta}|^{-2},$$

where the symbols are as for (35) with B_0, ϕ replaced by B_1, ϕ_1 as for (16).[4]

A_0+A_1 Mode

For a plane wave polarized perpendicular to the axis of a perfectly conducting (or ferromagnetic) cylinder, or for a rigid cylinder or one of large impedance in acoustics, and $Ka\to 0$, we retain only A_0 and A_1. Thus we obtain

$$\psi=\Delta(\mathfrak{N})H_0\{[A_0(1-2A_1X)-2A_1\sin\theta\sin\alpha(1-A_0X)$$
$$+2A_0A_1(\sin\theta-\sin\alpha)Y][(1-A_0X)(1-2A_1X)$$
$$-2A_0A_1Y^2]^{-1}+2A_1\cos\theta\cos\alpha\}, \quad (38)$$

which is the sum of $\psi^1=H_0\Delta(\mathfrak{N})[A_0+2A_1\cos(\theta+\alpha)]$ and the terms of (15). (Y is obtained from $X\propto T_-+T_+$ by changing the sign of T_+.) For $Kb\sin\alpha=2p'\pi$, Y vanishes and (38) can be factored further; we restrict ourselves to this case for simplicity.

For perfectly conducting (or rigid) cylinders ($A_0=-A_1=-i\pi x^2/4$) and $Kb\sin\alpha=2p'\pi$, (38) reduces simply to

$$\psi=\Delta(\mathfrak{N})H_0A_0\{(1+Be^{i\beta})^{-1}$$
$$+2\sin\theta\sin\alpha(1-2Be^{i\beta})^{-1}-2\cos\theta\cos\alpha\}, \quad (39)$$

where $B=tx^2(\pi/2Kb)^{\frac{1}{2}}$, $\beta=\tau+\pi/4$. For α or θ small, the maximal effects give rise to intensity minima, essentially as for the parallel polarization component for perfectly conducting cylinders, although these must now contend with the background factor $-2\cos\theta\cos\alpha$. For $\alpha=0$, $\theta=\pi/2$, or vice versa, (39) is formally identical with (17). For a given grating and λ, however, the effects are much less for this case than for the previous logarithmic coefficient, while the contrasts and intensities of the bands vary with λ as they did for the previous case of the perfect dielectric. Thus $B\propto\lambda^{-\frac{1}{2}}$ increases rapidly with decreasing λ, so that the most pronounced departures from single scattering theory occur for the shorter wavelengths, although these are not necessarily the darkest bands since the intensity ($\propto\lambda^{-3}$) increases as λ decreases.

For α and $\theta=|\theta|$ large, the maximal effects give rise to intensity maxima so that the effects are of opposite type to those that occurred for the parallel component. For α and $|\theta|=-\theta$ large, the intensity for certain values of δ may approach zero, the middle term canceling the first to yield much more pronounced dark bands than for the parallel component. For arbitrary α and θ, the components of ψ compete to determine the character of the band, and its structure can be determined only by numerical calculations. Thus for $\sin\alpha=1/4$ the band at $\sin\theta=0$, $g=+1$, is dark while the bands for $\theta=|\theta|$, $q=2$ to 5, are bright, the intensities increasing with increasing θ; however, for $\theta=-|\theta|$, $q=0$ to -3, the bands are initially dark but develop into "double bands" having both a pronounced maximum and minimum.

For the forward scattered wave or central maximum, $\theta=-\alpha$, the curly bracket of (39) yields

$$\{-1-Be^{i\beta}[(1+Be^{i\beta})^{-1}+4\sin^2\alpha(1-2Be^{i\beta})^{-1}]\}$$

so that for $\beta\approx 0$ the intensity $|\psi|^2$ is greater than its single scattered value for all α. For the backscattered wave $\theta=\pi-\alpha$, however, the curly bracket reduces to $\{3-Be^{i\beta}[(1+Be^{i\beta})^{-1}-4\sin^2\alpha(1-2Be^{i\beta})^{-1}]\}$ so that for $\beta\approx 0$ the intensity is less than its single scattered value for α small and greater for α large. The structure of these bands is in general much more complicated than for the A_0 mode, the band width and other features depending on both angular parameters.

The total scattered intensity and the interference term are given by

$$|\psi|^2=D^2\{4\cos^2\alpha\cos^2\theta+4R_1\sin^2\theta\sin^2\alpha$$
$$+R_0+4R_1\rho\sin\theta\sin\alpha$$
$$-4\cos\theta\cos\alpha[R_0(1+B\cos\beta)$$
$$+2\sin\theta\sin\alpha R_1(1-2B\cos\beta)]\}, \quad (40)$$
$$f=D\{2\cos\theta\cos\alpha\cos q-R_0(\cos q+B\cos q')$$
$$-2\sin\theta\sin\alpha R_1(\cos q-2B\cos q')\},$$

where

$$R_0=(1+2B\cos\beta+B^2)^{-1},\ R_1=(1-4B\cos\beta+4B^2)^{-1},$$

$\rho=(1-B\cos\beta-2B^2)$; the remaining symbols are as for (35) with $\phi=0$, $B_0=\pi x^2/4$.

** k_e = dielectric constant.

We note that the previous terminology, which was developed for the case where only one coefficient of the single cylinder was retained, is on the whole applicable to the present case provided that we retain the notion of a "mode hypothesis," i.e., the maximal effects for each mode occur for the wavelengths which insure optimal fulfillment of the conditions that each order of scattering of the mode is a maximum and that successive orders of scattering of the mode are either in or out of phase. For the perfect conductor, the maximal wavelengths (as well as those for the local extrema) are essentially the same for each mode (although the effects for each are of opposite character), so that the resultant peak effects should occur in their immediate vicinity and somewhat closer to those of the dominant mode. This is plausible for the bright peaks, but the minima of the dark bands may be shifted appreciably for θ and α on opposite sides of the surface normal, for which case the darkest parts of the bands may be determined by annulment between the modes rather than by the phase relations of the orders of scattering. Because of this "mode competition," the shapes of the bands may also differ appreciably from those of a single mode, and the width of the peaks and the strengths of the local extrema become functions of the angular parameters. More generally, when the phase difference of the two coefficients is not π, the effects of mode competition should be less pronounced but the displacement of the resultant maximal wavelengths greater.

IV. THE REFLECTION GRATING OF BOSSES ON A PLANE

Consider a plane wave incident on a reflection grating of semicylindrical bosses on a perfectly conducting infinite plane as at the right in Fig. 9. Applying the image technique discussed in Sec. 2,[4] we treat the problem of two plane waves, which are images in the removed plane, simultaneously incident on the grating of parallel cylinders.

For the component polarized perpendicular to the cylinders' axes or parallel to the plane of incidence (or for the acoustic case of a rigid plane), the total incident wave for the image problem is $\psi_i=\psi_i(\theta)+\psi_i(\pi-\theta)=\exp[iKr\cos(\theta+\alpha)+\exp[-iKr\cos(\theta-\alpha)]$. The total scattered wave is then $\psi''=\psi(\theta)+\psi(\pi-\theta)$ [where $\psi(\theta)$ is as in (1), (14), etc., while $\psi(\pi-\theta)$ has θ replaced by $\pi-\theta$], and the total reflected wave for the boss problem is $\psi_r''=\psi_i(\theta)+\psi''$; we regard $\psi_i(\pi-\theta)$, the wave on the right in Fig. 9, as the incident wave for the boss problem and $\psi_i(\theta)$ as the specularly reflected component. Thus in terms of Z as in (13) we obtain

$$\begin{aligned}\psi''&=2H_0\Delta(\mathfrak{N})(C_++Z),\\ C_+&=\sum\epsilon_nA_n\cos n(\alpha+\pi/2)\cos n(\theta-\pi/2).\end{aligned}\tag{41}$$

For the component polarized parallel to the cylinders' axes or perpendicular to the plane of incidence (or for the acoustic case of a plane of zero impedance), the total incident wave is $\psi_i(\theta)-\psi_i(\pi-\theta)$, the total scattered wave is $\psi'=\psi(\theta)-\psi(\pi-\theta)$, and the total reflected wave $\psi_r'=\psi_i(\theta)+\psi'$. Thus from (14) we obtain

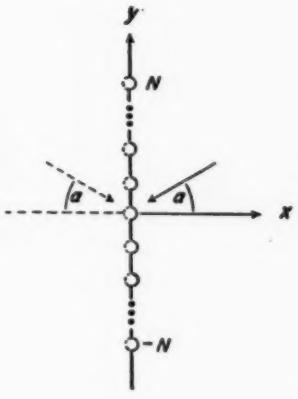


FIG. 9. Plane wave incident on reflection grating of $\mathfrak{N}$ parallel semicylindrical bosses on a perfectly conducting, or rigid, infinite plane. The image technique of solution is indicated.

$$\begin{aligned}\psi'&=2H_0\Delta(\mathfrak{N})C_-,\\ C_-&=-2\sum A_n\sin n(\alpha+\pi/2)\sin n(\theta-\pi/2),\end{aligned}\tag{42}$$

or the single scattered value obtained previously;[2,13] the multiple scattered terms vanish since they are symmetrical to the grating for $Kb\gg1$.

These expressions indicate that multiple scattering should primarily effect the component polarized perpendicular to the elements (which agrees with references 5–8). Thus if the bosses are dielectric and $x'\ll1$, the A_1 mode, we obtain

$$\psi''=2\psi_0,\quad \psi'=0,\tag{43}$$

where ψ_0 is as in (36). The function ψ'' also applies for the acoustic case of gaseous bosses on a rigid plane. If A_0 and A_1 are retained,

$$\psi''=2\psi_0,\quad \psi'=4A_1'H_0\Delta(\mathfrak{N})\cos\alpha\cos\theta,\tag{44}$$

where ψ_0 is the function of (38), (39) without the $\cos\theta\cos\alpha$ term. For the perfect conductor ($A_1'=-i\pi x^2/4=A_0$) for $A=2p'\pi$ we obtain

$$\begin{aligned}\psi''&=2\Delta(\mathfrak{N})H_0A_0[(1+Be^{i\beta})^{-1}+2\sin\theta\sin\alpha(1-2Be^{i\beta})^{-1}],\\ \psi'&=4\Delta(\mathfrak{N})H_0A_0\cos\theta\cos\alpha\end{aligned}\tag{44'}$$

[see Eq. (39)]; ψ'' and ψ' also apply for a rigid and open surface, respectively.

Hence the scattered component polarized perpendicular to the bosses will exhibit the same sort of bright and dark bands for particular values of Kb as appeared previously for both polarization components for the grating of cylinders. We will shortly discuss ψ'' of (44') in detail; for the moment we will consider ψ', which should also show these effects even if to a lower order of magnitude.

Proceeding to the next order of magnitude in Kb for ψ' we obtain from (3)

$$\psi'=4A_1'H_0\Delta(\mathfrak{N})\cos\theta\cos\alpha/(1-A_1'X'),\tag{45}$$

[13] V. Twersky, J. Appl. Phys. **22**, 825 (1951); J. Acoust. Soc. Am. **23**, 336 (1951).

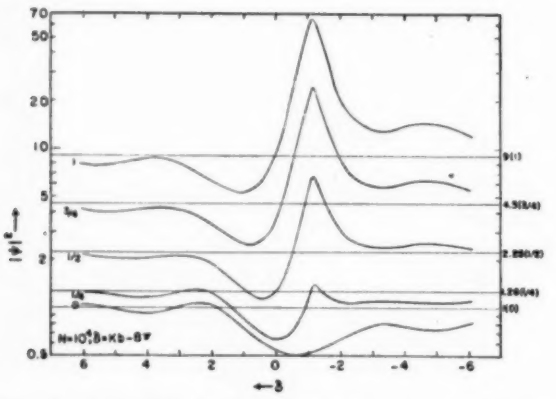


FIG. 10. Figures 10–19 are plots of the scattered intensity for the component polarized perpendicular to the elements of the perfectly conducting reflection grating (or a rigid reflection grating for acoustics); computations for the curves are accurate to about 5 percent except for $|\psi|^2 \to 0$. The normalization for all curves is $|\psi^1|=1$ for $\theta=\alpha=0$ (the backscattered wave for normal incidence) and all are for $N=10^4$; the numbers on the right scale are the single scattered values. The present graph shows the backscattered intensity for the vicinities of $\sin\theta=\sin\alpha=0$, 1/4, 1/2, 3/4, 1 for $\lambda=b/4$. Unity on the δ scale is 0.01 degrees for all graphs.

where

$$X'=2\sum[H_0(\sigma Kb)+H_2(\sigma Kb)]\cos(\sigma Kb\sin\alpha)$$
$$\approx(4i/Kb)\sum H_0(\sigma)\cos\sigma A/\sigma.$$

Hence for $A_1'=A_0$,

$$\psi'=4\Delta(\mathfrak{R})H_0A_0\cos\theta\cos\alpha[1+(2/Kb)B'e^{i\beta'}]^{-1}, \quad (45')$$

where $B'=t'x^2(\pi/2Kb)^{\frac{1}{2}}$, $\beta'=\tau'-\pi/4$, and $t'e^{i\tau'}=T'=\sum\sigma^{-\frac{1}{2}}e^{i\sigma Kb}\cos(\sigma Kb\sin\alpha)$. Multiple scattering effects should therefore also be observed for parallel polarization for the reflection grating, particularly for the longer wavelengths of the size of the spacing; departures from single scattering theory, however, should be in general much less pronounced than for the perpendicular component.

To obtain some idea of the relative magnitudes of the effects for the two components, we note that $\sum\sigma^{-\frac{1}{2}}\approx 2.3$, 2.5 for $N=10^2$, 10^4, respectively, as compared with $\sum\sigma^{-\frac{1}{2}}\approx 18.6$, 199 for those two values of N. Assuming that $Kb\approx 10$, β and $\beta'\approx 0$, and $B\approx 0.8$ so that $R=(1+B)^{-2}\approx 0.31$, yields $R'=(1+2B'/Kb)^{-2}\approx 0.96$, 0.996 for $N=10^2$, 10^4, and the effects are negligible for the parallel component. We note, however, that for other values of the parameters, R' may differ appreciably from unity, and also that for certain values of α and θ the parallel effects may be more pronounced than the perpendicular, the maximal effects yielding a dark band for the former, and a dark or bright band for the latter. An interesting feature of these "parallel anomalies" for the reflection grating is that since $\beta'=\tau'-\pi/4$, τ_A' and consequently δ_A, as for (25), must be positive quantities. Hence the darkest part of the band is on the short wavelength side of $\lambda_0=b/p$, or the peak effects are shifted to the blue rather than the red side of λ_0. Faint parallel anomalies were observed by Wood[5] with reflection gratings yielding pronounced perpendicular anomalies, but these were not investigated. More recently a

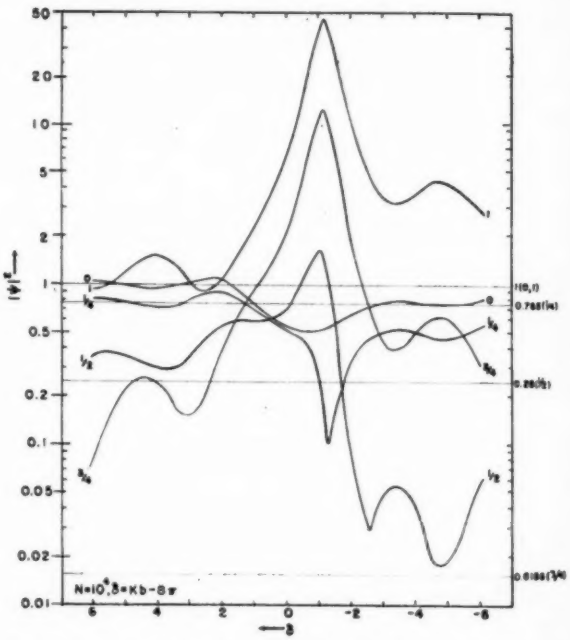


FIG. 12. Plots of the "forward scattered" intensity, or central maximum, for the vicinities of $-\sin\theta=\sin\alpha=0$, 1/4, 1/2, 3/4, 1 for $\lambda=b/4$.

detailed investigation of these anomalies was carried out by Palmer,[8] but the results have not yet been published.†† It should be noted that Rayleigh's approach,[9] to the order of approximation to which it has been carried, does not predict the parallel effects.

The total scattered intensity for the perpendicular component of (44') is

$$|\psi''|^2=4D^2(R_0+4R_1\sin^2\theta\sin^2\alpha+4R_0R_1\rho\sin\theta\sin\alpha), \quad (46)$$

where R_0, R_1, and ρ were defined for (40). The essential features of this expression were discussed with reference

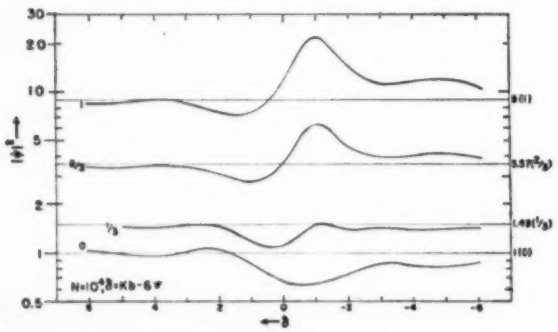


FIG. 11. Backscattered intensity for $\sin\theta=\sin\alpha=0$, 1/3, 2/3, 1 for $\lambda=b/3$.

†† *Note added in proof:* C. H. Palmer, J. Opt. Soc. Am. **42**, 269 (1952). Palmer found that the parallel effects were more pronounced than the perpendicular for some cases, and that in general the two sets for a particular grating were related only in that their sharp boundaries occurred for the values of (7').

to (39′) for the transmission grating (which had an additional term in $\cos\theta\cos\alpha$); it differs significantly from that obtained for isotropic scatterers in that the character of the bands depend markedly on α and θ.

Figures 10–19 are of $|\psi|^2$ *vs* $\delta = Kb - 2p\pi$ for $N = 10^4$; the even numbered ones are for $\lambda = b/4$ while the odd are corresponding plots for $\lambda = b/3$. It is assumed in these curves, as in the previous, that θ is varied with λ to correspond to a principal maximum of $\Delta(\mathfrak{N})$ for a given value of α so that $|\psi^1|^2$ can be considered constant for each curve; $(1+2\sin\alpha\sin\theta)^2$ is insensitive to the small variations of θ this requires for δ small. The normalization for all curves is $2B_{\max} = 0.9$ for $\lambda = b/4$ so that the ratio of radius to spacing is 3.8×10^{-3}; the maximum value of $2B$ for $\lambda = b/3$ is 0.585. Unity on the logarithmic intensity scales for $\lambda = b/4$ is $4D^2 = \mathfrak{N}^2x^4(\pi/2Kr)$, while that for $\lambda = b/3$ is 0.422 times this quantity; values of $|\psi^1|^2$ are on the right scale. Note that since

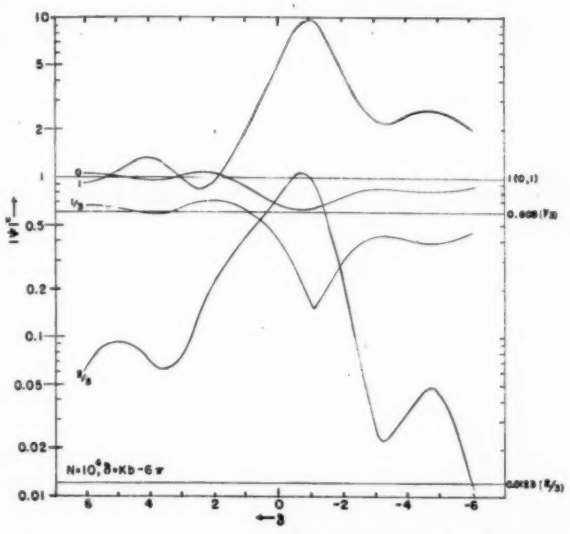


FIG. 13. Forward scattered intensity for $-\sin\theta = \sin\alpha = 0$, 1/3, 2/3, 1 for $\lambda = b/3$.

$\Delta_3 = 16\Delta_4/9$, the bands for $\lambda = b/3$ are about 1.78 times the width of those for $\lambda = b/4$ on a wavelength scale. Unity on the δ-scale is 0.01 degrees for all groups.

Figures 10 and 11, which show the backscattered intensity near $\theta = \alpha = \sin^{-1}(p'/p)$, $p' = 0, 1, \cdots p$, indicate that the bands that are dark for normal incidence grow into bright bands with increasing α and θ, the maximum shifting to the long wavelength side of the minimum. The greatest departures from single scattering theory, $R_\wedge \equiv R_{\max}$, $R_\vee{}^{-1} \equiv R_{\min}{}^{-1} (R = |\psi/\psi^1|^2)$, and the contrasts, $R_\wedge/R_\vee$, increase in general with increasing α (curves marked 1/4 and 1/3 are exceptions) and are greater for the shorter wavelength. The darkest bands, however, occur for the longer wavelength. Figures 12 and 13, which show the "forward scattered" intensity, or central maximum, near $\theta = -\alpha = -\sin^{-1}(p'/p)$ indicate that with increasing α the bands initially become more pronounced dark bands but that

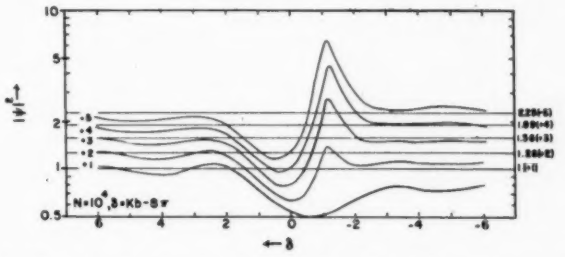


FIG. 14. Plots for the vicinities of the principal maxima in the positive spectral orders ($q = 1$ to 5) for $\sin\alpha = 1/4$, $\lambda = b/4$.

as α increases further they turn into bright bands, the maximum (which has broadened a great deal) shifting to the long wavelength side of the minimum. In this process bands may appear which have a pronounced maximum but "tail off" (i.e., which have their most pronounced local maxima) on their short wavelength side as in the 2/3 curve, while other bands may be "double," possessing both a marked maximum and minimum as in the 1/2 curve. The maximal departure and the contrast increase initially with increasing α but fall off as α increases still further; these are in general greater for the shorter wavelength. The 3/4 curve, one of the exceptions to this as far as $R^{-1}{}_\vee$ is concerned, is in this region of δ entirely above its single scattered value. The value of $R_\wedge$ for the 3/4 curve is the largest intensification we have encountered.

Figures 14 and 15 show the scattered intensity for the positive orders for $\sin\alpha = 1/p$, and Figs. 16 and 17 show the central maximum and negative orders; the positive orders lie on the side of the central maximum towards the source, or counter clockwise from the direction of the image wave (the direction of the central maximum) in Fig. 9. The behavior of the curves for the positive orders with increasing q is essentially as for the backscattered waves of Figs. 10 and 11 with increasing α, although the growth is more gradual and the effects not so pronounced; the curves for the central maximum and the negative orders are all dark bands whose complexity increases with increasing q, one of the local minima finally turning into the darkest part of the band. It can be seen that $R_\wedge$ for the negative orders is in general greater for the shorter wavelength, while $R^{-1}{}_\vee$ and $R_\wedge/R_\vee$ are greater for the longer wavelength, i.e., the more pronounced intensification occurs for $\lambda = b/4$, while the more pronounced reduction (as well as the darker bands) and bands of greater contrast occur for $\lambda = b/3$ for the negative orders. The value of $R_\vee$ for $\lambda = b/3$, $q = -2$, is the largest

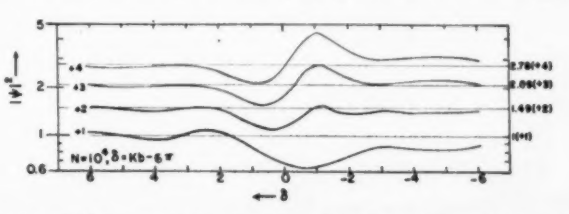


FIG. 15. Positive spectral orders ($q = 1$ to 4), for $\sin\alpha = 1/3$, $\lambda = b/3$.

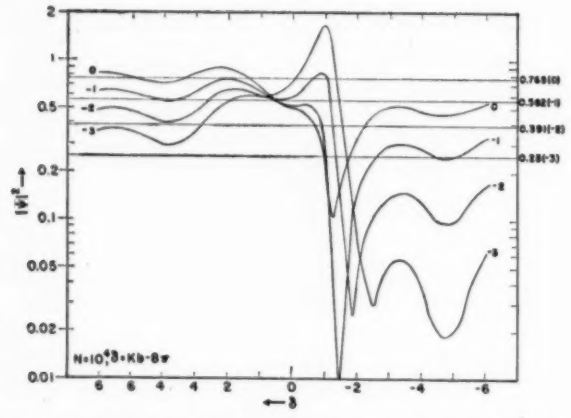


FIG. 16. The central maximum and negative spectral orders ($q=0$ to -3) for $\sin\alpha=1/4$, $\lambda=b/4$.

reduction we have encountered; unlike the case of isotropic scatterers $|\psi|^2$ may equal zero, e.g., when $\sin\theta\sin\alpha=-(1-2B)/2(1+B)$, provided that θ and α are permissible values. The value of $R_\wedge/R_\vee$ for this curve is the greatest contrast yet encountered and this band is also the darkest.

Figures 18 and 19 show the scattered intensity for $\sin\alpha=p'/p$, and $q=\pm1$; each curve corresponds to two sets of values because of the reciprocity of θ and α. Increasing α from 0 to $\pi/2$ and observing the positive order, we go through curves *a*, *a*, *b*, *c*, (*d*), while for the negative order we go through *a*, *b*, *c*, (*d*). Both sets of orders start out initially as dark bands and finish as bright, but the effects for $q=+1$ "lag" those for $q=-1$, the same curves appearing for $q=+1$ for a larger value of α. Comparison of Fig. 18 with Fig. 12 and Fig. 19 with Fig. 13 indicates that for a given value of α the bands for the central maximum and one of the adjacent orders may "tail off" in a direction opposite to the band of the remaining adjacent order, while for other values the bands for the central maximum and one adjacent order may complement the character of the band of the remaining order. Thus $\sin\alpha=1/2$, $q=0$ and -1 show maxima while $q=+1$ is a pronounced minimum; $\sin\alpha=3/4$, $q=0$ and -1 yield pronounced bright bands completely above their single scattered values for this region of δ, while for $q=+1$ the maximum tails off in the opposite direction; $\sin\alpha=2/3$, $q=0$ and -1 show marked maxima tailing off in opposite directions, while $q=+1$ is a pronounced minimum; and finally for $\sin\alpha=1$, $q=0$ and $+1$ show the same sort of bright bands for both sets. (This sort of behavior has been observed by Wood[5] and Strong.[6]) The intensification $R_\wedge$ increases with α for both wavelengths, the effect being greater for the shorter; the reduction, $R^{-1}{}_\vee$, has its largest values for the *b* curves where the effect is greater for the longer wavelength, although for the other curves the effects are greater for the shorter; while the contrast has its greatest value for the *c* curve for the longer wavelength and for the *b* curve for the shorter, the values being greater for the longer except for curve *b*. We note that for $q=+1$ the bands for $\sin\alpha=0$, $1/4$ are identical, while the minimum for $\sin\alpha=1/2$ is much narrower; similarly for $q=+1$ $\sin\alpha=0$, $1/3$ and $\sin\alpha=2/3$. [This sort of behavior seems indicated in Fig. 1.3, Wood,[5] (1935).]

These curves were computed for $N=10^4$, but as can be seen from comparison of Figs. 5–7, we would obtain essentially the same results for $N=10^3$ and 10^2 except that the values on the δ-scale would be multiplied by 10 and 100, respectively.

For simplicity we have considered only the contribution scattered by the grating in the above. More generally, however, we must consider the total reflected wave, or the sum of the scattered component and the specularly reflected image wave. Thus for an incident wave as in Fig. 9, the total reflected fields and intensities are as in (21),[4] (22).[4] The required interference terms are

$$\begin{aligned} f''&=-2D[R_0(\cos q+B\cos q')\\ &\quad+2R_1\sin\theta\sin\alpha(\cos q-2B\cos q')],\\ f'&=-4D\cos\theta\cos\alpha\cos q, \end{aligned} \tag{47}$$

where the notation is as for (40). The discrepancies that may arise between the curves here presented and those obtained experimentally, and their causes are essentially as for the transmission grating as discussed previously.

V. EXTENSION OF THE THEORY TO GRATINGS WITH ELEMENTS OTHER THAN CYLINDERS

In this paper we have presented a multiple scattering theory of the grating of cylinders. A formal solution to the mathematical problem was obtained, approximate solutions were analyzed in detail, and the results were compared with experiment. The departures from elementary theory that the present theory predicts were correlated with the experimental "grating anomalies," and their presence interpreted by simple physical

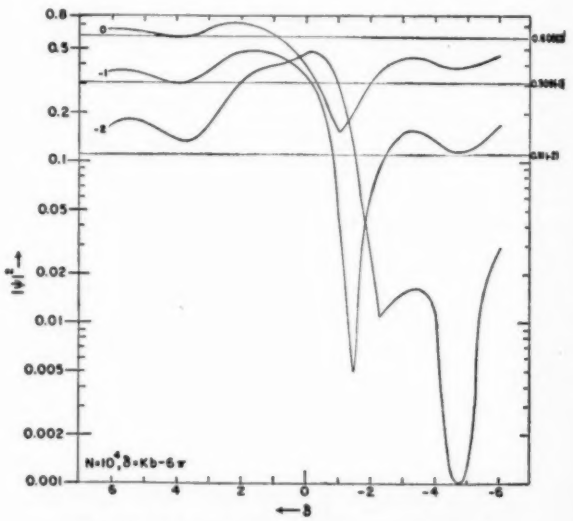


FIG. 17. Central maximum and negative orders ($q=0$ to -2) for $\sin\alpha=1/3$, $\lambda=b/3$.

arguments. Only gratings of circular cylinders or semicylindrical bosses on a plane were considered, but it is evident that the theory can be extended to transmission gratings whose elements are elliptical cylinders, as well as to the corresponding reflection gratings. For the restrictions for which the solution in closed form obtains, the appropriate representations for certain cases can be written down by inspection; similarly, for the limiting cases of very small elements, analogous mode approximations in which only one or two scattering coefficients of the single element are retained can be found.

The significance of the scattering coefficient of the single element in controlling the form of the intensity curves is evident from the preceding sections. Although the physical condition that each order of scattering of the grating be a maximum is independent of the elements, being identically (7) for all gratings, fulfillment of the conditions that successive orders be in or out of phase is dependent on the phase of the scattering coefficient (or the material and polarization component). Similarly the contrast and the over-all intensity of the bands as functions of λ depend on the form of the modulus of the coefficient. Thus for the cases treated in Figs. 5–8, the bands were dark or bright as the phase was 0 or π, while the contrast and over-all intensity increased with λ for the logarithmic coefficient and decreased with increasing λ for the coefficient in λ^{-2}. Similarly for two modes, where the additional factor of mode competition exists, the bands were bright or dark depending on the dominant mode as determined by the angular parameters and the phases of the coefficients, while the contrast and over-all intensity were in general

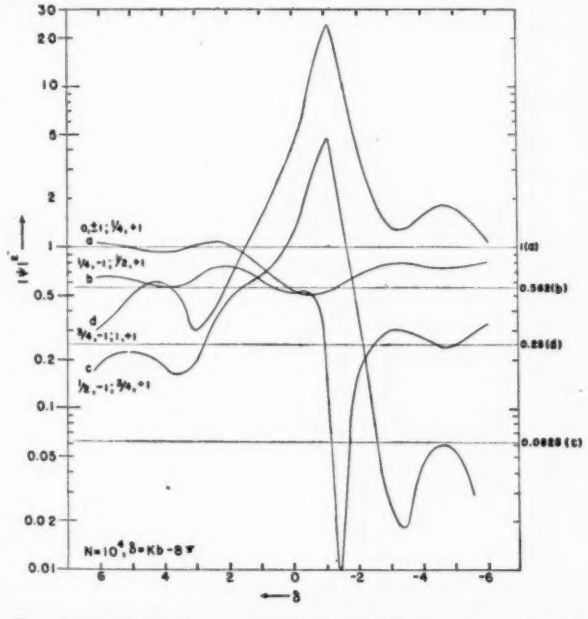


FIG. 18. Plots for $\sin\alpha=0$, 1/4, 1/2, 3/4, 1 and $q=\pm1$, the first spectral orders on each side of the central maximum, for $\lambda=b/4$; e.g., curve b corresponds both to $\sin\alpha=1/4$, $q=-1$ and $\sin\alpha=1/2$, $q=+1$ (because of reciprocity in θ and α).

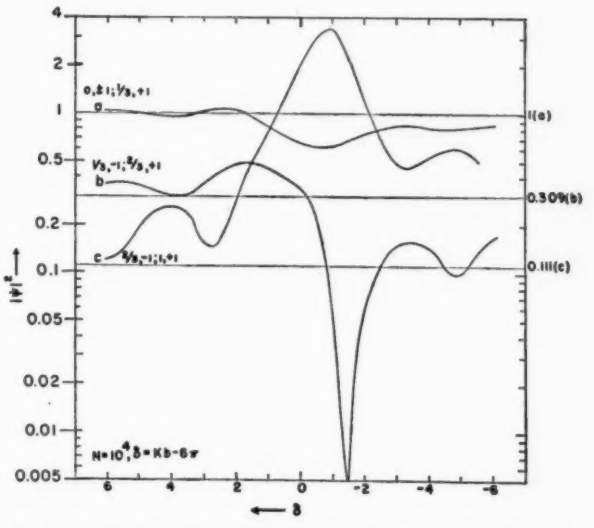


FIG. 19. Plots for $\sin\alpha=0$, 1/3, 2/3, 1 and $q=\pm1$ for $\lambda=b/3$.

larger for the shorter wavelength because of the λ^{-2} behavior of both coefficients.

We can extend the theory to gratings with elements other than cylinders by employing the approximate solutions for a perfectly conducting single strip and a single slit in a perfectly conducting plane which were derived by Rayleigh.[14] For the strip and perpendicular polarization, or for the slit and parallel polarization, he obtained essentially the A_1 mode with A_1 as for perpendicular polarization for a perfectly conducting cylinder whose diameter equalled the slit or strip width. Similarly for the strip and parallel polarization, or the slit and perpendicular polarization, he obtained the A_0 mode with A_0 as for parallel polarization for a perfectly conducting cylinder. The extension of multiple scattering theory to these elements can be carried out by substituting these values into the expressions we have derived, a grating of narrow strips corresponding to our transmission grating of cylinders, and that of slits to our reflection grating.

It is tempting to extend Babinet's principle, as illustrated by the above, to the analogous pair of the semicylindrical boss and a semicylindrical depression; the second of these is a difficult problem mathematically and little has been accomplished towards its solution. Proceeding formally, we replace the coefficient amplitudes in λ^{-2} for the perpendicular polarization component by the logarithmic amplitude that appeared for the parallel component, and vice versa, to obtain expressions indicating essentially the same sort of bands occurring for the same wavelengths and angles as for the bosses; they differ only in that their contrasts and relative intensities *increase* with increasing λ rather than decrease as for the bosses. If this procedure could be justified, we would obtain a more suitable model for the reflection grating with grooves narrow compared

[14] Rayleigh, Phil. Mag. 43, 259 (1897).

to wavelength. We have brought this point up to indicate the basis of the only marked discrepancy between our theoretical results for the reflection grating of bosses and the available experimental results for narrow-grooved gratings: the experimental results show the same increase of contrast with increasing λ that is deduced from the expressions obtained by extending Babinet's principle to the boss to obtain the solution for the depression.

The experiments of Wood,[5] Strong,[6] Ingersoll,[7] and Palmer[8] mentioned previously were performed with essentially two types of reflection gratings. In some cases the anomalies could be traced to the presence of fine ridges or proturberances bordering the usual grating grooves (as was deduced on the disappearance of the anomalies from the spectra by gently rubbing the grating), while in others it was surmised that the effects were due to the very narrow grooves themselves. Our results in general agree with these experiments, particularly with those for the "ridged" gratings for which Wood[5] (1935) also observed fainter anomalies polarized parallel to the elements. The intensity curves we have sketched are more complicated than those obtained experimentally, the discrepancies no doubt arising from the approximations employed, the presence of the incident and reflected plane waves, and the effects of the particular groove form; the last of these is perhaps the most significant factor.

The asymmetry of the grating groove results in the multiple excitations traveling in one direction along the grating differing in magnitude and phase from those traveling in the other; this could account for the absence or disappearance of certain of the branches in Wood's spectrograms and also for the fact that in many cases the intensity of these branches differs little from the intensity at the intersections where $A=2p'\pi$. [These cases, where only the waves going up or down the grating exist, correspond to the vanishing of either S_- or S_+ in (14).] Similarly, exact correspondence is not expected for the ridged gratings since they actually comprise two superposed gratings, one of fine ridges which multiple scatters, and one of wide grooves which only singly scatters; for the second of these, presumably the large width of the grooves compared to wavelength precludes satisfying the optimal phase relations required for pronounced multiple scattering effects. Such gratings therefore possess additional background Fraunhofer spectra which mask some of the effects of the ridges.

The writer is grateful to Professor Fritz Reiche for suggesting the research and to Professor Joseph B. Keller, Professor Wilhelm Magnus, and Dr. Jerry Shmoys for many fruitful discussions.

APPENDIX‡‡

Derivation of Z of (13)

From an examination of Z of (9) it follows that it can be written as $Z=C_+'C_-''K_{11}+C_+'C_+''K_{12}+C_-'C_-''K_{21}+C_-'C_+''K_{22}$, where the K's are power series in S_-, S_+ with coefficients which are multiples of products of powers of C^0, C^π.

From the general expression for Z we find that the first terms of K_{11} are $K_{11}=S_-+S_-^2C^0+S_-^3(C^0)^2+S_-^2S_+C^0C^\pi+\cdots$. These terms arise from the following process: we consider a product of $(m+1)$ brackets

$$[S_-+(-1)^{n+n'}S_+][S_-+(-1)^{n'+n''}S_+]\cdots[S_-+S_+\exp\{i\pi[n^{(m)}+n^{(m+1)}]\}].$$

Any term arising from the multiplication of these brackets which is involved in K_{11} must necessarily be such that the first and the last bracket contribute the factor S_-. This follows from the assumption that K_{11} involves the terms to which the C_n', C_n'' contribute a factor which is a product of $C_+'C_-''$. Now let us denote by D_m the terms of K_{11} which are of degree $m+1$ in S_-, S_+. We split D_m into a sum, $D_m=u_m+v_m$, where u_m involves all the terms of D_m which are obtained by choosing S_- as the contribution of the mth (the second, but last) factor. Similarly, v_m denotes the terms which arise if we take $S_+\exp\{i\pi[n^{(m)}+n^{(m+1)}]\}$ as the contribution of the mth bracket. Then u_{m+1}, v_{m+1} are connected with u_m, v_m by a linear substitution, the matrix of which,

$$M=\begin{pmatrix} S_-C^0 & S_-C^0 \\ S_+((C^\pi)^2/C^0) & S_+C^0 \end{pmatrix},$$

does not depend on m (see appendix, footnote †, for details).

Now since $I+M+M^2+M^3+\cdots=(I-M)^{-1}$, where I denotes the identity matrix, and since $u_1=S_-^2C^0$, $v_1=0$, we find after an elementary calculation

$$K_{11}=D_0+\sum_{m=1}^{\infty}(u_m+v_m)=S_-(1-S_+C^0)\times\{1-(S_-+S_+)C^0+S_-S_+[(C^0)^2-(C^\pi)^2]\}^{-1}.$$

The coefficients K_{12}, K_{21}, K_{22} can be computed by the same method. The denominator, which is the determinant of $I-M$, will always be the same.

‡‡ This appendix was prepared by Professor Wilhelm Magnus.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 10 OCTOBER, 1952

On the Use of the Joule-Thomson Effect to Measure the Work Equivalent of Heat. Thermodynamic Theory

JOHN R. ROEBUCK
The University of Wisconsin, Madison, Wisconsin
(Received March 19, 1952)

The Joule-Thomson porous plug experiment is described, its quantitative aspects indicated, and the constant enthalpy condition set up. The experimental plan for the proposed measurement is described and a short analysis of the physical situation in the fluid given. The standard equation for the Joule-Thomson coefficient is applied to the situation and details are given of how this equation may be integrated for two special physical situations. The information necessary for these integrations is discussed and physical means described for obtaining it. A numerical integration may be used for obtaining the work equivalent of heat (J) with a high precision corresponding to the high precision with which these various data may be measured.

THE porous plug experiment of Joule and Thomson has long offered an attractive method for measuring the mechanical equivalent of heat. This was noted by Maxwell in his little book on heat, and probably has been realized by many others. The author[1,2] spent his research time between 1908 and 1912 on a first attempt to work out the details of a satisfactory theory and an effective experimental procedure. This effort is continued in the work reported here.

In this classical experiment a fluid is driven along a passage containing a porous obstruction in which it suffers a drop in pressure and usually a change in temperature. Such steady-flow situations offer great advantages for precise measurement, in that states transient in an element of the fluid become fixed in space and unvarying in time, and so may be measured at leisure.

Elementary textbooks on thermodynamics derive the equation,

$$u_1+p_1v_1=u_2+p_2v_2=h=\text{enthalpy},$$

where the derivation assumes that (1) the experiment is carried out without any exchange of heat by the fluid with its environment, (2) the only forms of energy involved are heat and work, and (3) the first law holds. That is to say, the process between the thermometers runs at constant enthalpy, and the space from the first thermometer through the plug to the second thermometer must be kept adiabatic. Moreover, the only pressure changes between the thermometers must be those across the plug itself.

Our experimental method[3] of making these measurements is to hold the T_1 and p_1, at entry to the plug, constant by suitable thermostat and barostat. The fluid escapes from the plug under control of a second barostat set at any one of a group of arbitrary pressures, p_i, between p_1 and p_2, and the T_i associated with each p_i is read. This group of pairs of T's and p's are plotted and a smooth curve drawn through them. We call this an "isenthalpic curve."

Since pressure and temperature are measurable only in equilibrium states, the irreversible turbulent process in the plug has been interrupted at any arbitrarily desired pressure, and the temperature of the resulting equilibrium state measured. These measured states may be brought as close together as desired, and in principle the points may be as numerous as desired.

We thus set up a series of equilibrium states separated by sections of the actual turbulent process. The state set up in each equilibrium is necessarily determined by what happens in the preceding turbulent section. It is apparent that nothing can happen in the turbulent section which will prevent the fluid from settling down to an equilibrium state with the same enthalpy as the preceding equilibrium states showed.

One thing which does happen in the turbulent section is an occurrence of gross turbulence. By this is meant some coordination in movement by large groups of molecules. This coordinated energy of the group is not appreciated fully as heat by a thermometer, the total heat of the region is lessened, and the temperature there falls. This phenomenon shows up very clearly when the volume passing the second thermometer is large and therefore the escape velocity is also large. The plug drop of temperature observed increases, and the point falls below the curve for smaller escape velocity. It might be noted that here the escape velocity is roughly parallel to the thermometer wall.

The plugs[4] we use are very fine-grained unglazed porcelain tubes with one hemispherical closed end. The higher the temperature of firing the porcelain the less the permeability of the tube wall. We select from our stock of tubes one with a permeability suited to the pressure drop and mass flow planned. The area exposed to normal flow is considerable and the wall thickness small. All these factors together result in a plug having a network of holes, very numerous, very small, tortuous, and irregular in size. Such passageways quickly degrade the coordinated movement to the random molecular motion we picture as heat.

The proportion of heat energy diverted at any time

[1] J. R. Roebuck, Phys. Rev. **2**, 79 (1913).
[2] J. R. Roebuck, Phys. Rev. **2**, 299 (1913).
[3] J. R. Roebuck, Proc. Am. Acad. **60**, 537 (1925).
[4] Burnett and Roebuck, Phys. Rev. **30**, 529 (1910).

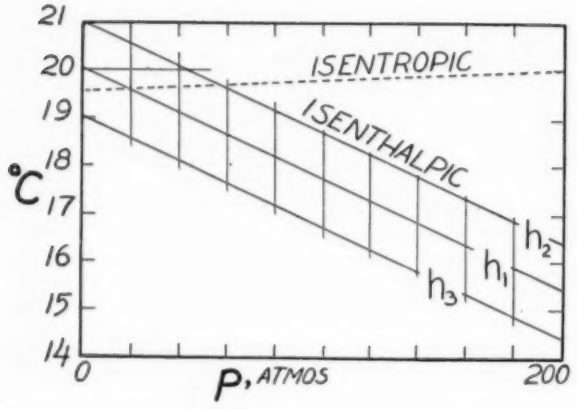


FIG. 1. Isenthalpic and isentropic expansion of water.

into this gross turbulence remains small. A relatively small adjustment serves to restore the normal temperature equilibrium. We have been unable to find evidence of such gross turbulence in the gas escaping from the plug wall. We have assured ourselves in the gas work that it did not persist to the second thermometer in sufficient amount to be detected in its reading. It seems reasonable to conclude that the state of the fluid flowing in the porcelain is not greatly different from its equilibrium state.

Incidentally one might note that these considerations do not apply to a nozzle or valve flow.

Application of the first and second laws to this situation leads directly to the well-known Eq. (1) for the Joule-Thomson porous plug coefficient μ, namely,

$$\mu C_p \equiv C_p\left(\frac{\partial T}{\partial p}\right)_h = T\left(\frac{\partial v}{\partial T}\right)_p - v, \qquad (1)$$

where $h=u+pv=$enthalpy.

Put C_p in heat units and the other quantities in mechanical units. Separate the quantities in $(\partial T/\partial p)_h$ and integrate formally; thus

$$\left[J\int_{T_1}^{T_2} C_p \partial T = \int_{p_1}^{p_1} T\left(\frac{\partial v}{\partial T}\right)_p \partial p - \int_{p_1}^{p_2} v\partial p\right]_h, \qquad (2)$$

where the constancy of h must be maintained in each of the three integrals. Or said differently, this requires that the integrals be taken along the above constant enthalpy curve. To carry out the detailed integration of Eq. (2), it is necessary to know the values of the coefficients C_p, T, $(\partial v/\partial T)_p$ and v along the isenthalpic curve, at least at the measured points.

The behavior of gases and liquids in porous plug experiments presents marked differences. The changes of volume of gases under the pressure changes used in these experiments are very great and the work done by the gas in expanding down the pressure gradient is correspondingly great. In our experiments this expansion takes place in the very small pores of our unglazed porcelain tube walls. The pressure gradient set up in the chain of pores accelerates the gas, but this mass kinetic energy is quickly degraded to heat in these small irregular passages, a process usually called friction. Therefore, effectively the work of nearly isothermal expansion is left in the gas as heat. The resulting change in temperature across the plug is usually not large and may be positive or negative. Liquids have a much greater density, a much greater volume heat capacity, and a much smaller change of volume with change of pressure. The work of expansion of the liquid down the pressure gradient in the plug is quite small, and the work done to drive unit volume through the plug is mainly occupied in raising, through gross turbulence and friction, the temperature of the liquid a small amount. The greater viscosity and lower velocity of escape of liquid under the same circumstances should make the persistence of gross turbulence in the fluid leaving the plug even less than for the gas.

Thus the isenthalpic curves for water have the character shown in Fig. 1. They slope up with falling pressure and have slight curvature. The different curves are characterized by the value of h. If the h_1 curve runs through the point 20°C and one atmosphere, then for present purposes we define C_p as 1 cal/g °C there. Also $C_p \equiv (\Delta h/\Delta t)_p$ by definition, so that

$$C_{p_1} = \frac{h_2 - h_3}{\Delta T} = \frac{\text{const}}{\Delta T} \quad \text{or} \quad C_p \cdot \Delta T = \text{const}.$$

Hence it is planned to measure ΔT intercepts at a series of measured pressures to that C_p may be calculated with precision along curve h_1. C_p will then be known as a function of T and p along the isenthalpic curve for use in the integration of Eqs. (2) and (6).

The isenthalpic curve itself gives $T=f(p)$ and the T values at the measured points are known.

The coefficient of expansion $1/v(\partial v/\partial T)_p$ and the specific volume v, for water, both as a function of p and T, have been measured carefully, and the values are available in the literature.

Thus we have the needed four coefficients as functions of p and T along the isenthalpic curve and are in a position to carry out the desired integration of Eq. (2).

A rough calculation shows that the first term on the right in Eq. (2) is only about 5 percent of the second. This is most fortunate as $(\partial v/\partial T)_p$ is much the least precisely measured of these coefficients while the specific volume is probably known the most precisely.

It should also be noted that the first term on the right of Eq. (1) vanishes when $(\partial v/\partial T)_p$ vanishes at the temperature of maximum density. The temperature of maximum density falls with rising pressure. To make the corresponding integral term in Eq. (2) vanish, it would become necessary to divide the range so that a negative and a positive part cancel each other. No direct experimental method for this seems available. The use of an unusual working temperature would com-

plicate the provision of data for the other integrating coefficients.

Another procedure is available for obtaining these integrals from somewhat different experimental data.

Maxwell's fourth equation,[5] namely,

$$\left(\frac{\partial s}{\partial p}\right)_T = -\left(\frac{\partial v}{\partial T}\right)_p,$$

when combined with the general relation[5] for the variables (p, T, s), namely,

$$\left(\frac{\partial p}{\partial T}\right)_s \left(\frac{\partial T}{\partial s}\right)_p \left(\frac{\partial s}{\partial p}\right)_T = -1,$$

yields

$$C_p\left(\frac{\partial T}{\partial p}\right)_s = T\left(\frac{\partial v}{\partial T}\right)_p, \tag{4}$$

which when substituted in Eq. (1) above gives

$$\left(\frac{\partial T}{\partial p}\right)_h = \left(\frac{\partial T}{\partial p}\right)_s - \frac{v}{C_p J}. \tag{5}$$

As with Eq. 1, put C_p in heat units and the other quantities in mechanical units. Separate the quantities in $(\partial T/\partial p)_h$ and integrate formally; thus

$$\left[\int_{p_1}^{T_2} \partial T = \int_{T_1}^{p_2} \left(\frac{\partial T}{\partial p}\right)_s \partial p - \frac{1}{J}\int_1^{p_2} \frac{v}{C_p}\partial p\right]_h. \tag{6}$$

The third term from the left may be integrated along the isenthalpic curve with the help of the knowledge of v and C_p discussed above. To integrate the second term it is necessary to know $(\partial T/\partial p)_s$ along the isenthalpic curve. This coefficient is the rate of adiabatic change of temperature with pressure.

To understand the part played by this coefficient, $(\partial T/\partial p)_s$, in the porous plug experiment, consider the physical process involved. The liquid water is driven through the pores of the plug by the water pump against the viscous friction in these small irregular passages, and the heat from this friction is carried forward with the water. As the pressure drops continuously along the flow, the liquid itself expands so as to keep the equilibrium between the volume and pressure. This volume expansion of the liquid takes place against the residual pressure, tending to maintain it, and the work energy required is drawn from the heat energy of the liquid. Such an expansion, maintaining the pv equilibrium, is a reversible one. Experimentally the expansion takes place under strictly adiabatic conditions. There is thus a reversible adiabatic (i.e., isentropic) expansion taking place in the liquid on its way through the plug and drawing its energy from the heat energy of the liquid. With water around 20°C, this expansion reduces the rise of temperature resulting from the viscous friction.

Water cools on reversible adiabatic expansion by about 0.3°C for 200 atmos drop, starting the drop with the water in the neighborhood of 20°C. One can obtain the data for an adiabatic (i.e., isentropic) curve by dropping the pressure on water stationary in a strong container, successively to a series of arbitrary values, and reading each associated resulting temperature (see Fig. 1). The water before each expansion must be at the same pressure and temperature. The slope of such a curve at a point is the value of $(\partial T/\partial p)_s$ there. Choose the series of points so that the resulting curve crosses the isenthalpic curve, and at the crossing point pick off the slope of the isentropic curve. Do this for a sufficient number of points along the isenthalpic curve. These values are those needed for the integration of the second term in Eq. (6) above.

Thus the three terms of Eq. (6) may be integrated and the value of J obtained.

We have thus at present two possible methods for obtaining J from Joule-Thomson data. The data is in part different in the two cases. It would seem at present that the process using Eq. (2) promises the more precise value of J.

Experimental work has already proved that it is possible to duplicate readings on the multiple U-tube mercury manometer, of the pressure resulting from a loaded rotating piston, to at least 1 part in 10^4, at 200 atmos. The next step, for which apparatus is already set up and tested qualitatively, is to measure the precision with which the change of temperature on full drop of pressure may be duplicated. For the integration to obtain J, several isenthalpic curves will be measured with high precision, for which apparatus is already prepared. It is proposed to publish an article on the mercury manometer next.

Help has been received from many people, in particular from Dr. Harold Osterberg, in the working out of this theory, which is hereby gratefully acknowledged.

Thanks are also due the Office of Naval Research for their continuing support of this project, and the Department of Physics for the freedom given me in the use of the Department facilities.

[5] M. W. Zemansky, *Heat and Thermodynamics* (McGraw-Hill Book Company, Inc., 1951), third edition, pp. 31 and 246.

On the Electrical Properties of Porous Semiconductors*

EUGENE B. HENSLEY†
Department of Physics, University of Missouri, Columbia, Missouri
(Received March 31, 1952)

It has been proposed that the conductivity of a porous semiconductor, such as oxide-coated cathode, may be influenced by the presence of an electron gas in the pores of the aggregate. In this paper calculations for the magnitude of this component of the conductivity are made on the basis of two simplified models approximating the geometry of a pore. The conditions under which such a component of the conductivity can appreciably modify the total conductivity are analyzed. It is further shown that such a porous semiconductor will possess two sources of thermoelectric emf. A simple theory for the thermoelectric power of the electron gas in a pore is developed, and the manner in which it will combine with the normal thermoelectric power of the crystals is shown.

I. INTRODUCTION

NUMEROUS attempts have been made to explain the electrical properties of oxide-coated cathodes.[1] In most recent years it has appeared as though most of the properties could be correlated on the hypothesis that the crystals of the cathode coatings were simple n-type semiconductors. Recently, however, Loosjes and Vink[2] have presented a theory which states that the conductivity of an active oxide cathode at high temperatures is a property of the electron gas in the interstices between the crystallites of the porous structure. In this paper we will examine the conditions under which a porous aggregate of crystals which are n-type semiconductors will have its conductivity influenced by the electron gas in the pores and will discuss further consequences of such a system.

There is little doubt that an electron gas exists in the pores of an oxide cathode since this is the equivalent to the space charge observed at the cathode surface. Also, because the electron affinity χ of the crystals is very probably positive, the density of these electrons will be considerably less than the density of electrons in the conduction band of the crystals. However, since the dimensions of the pores in the cathode are believed to be of the order of 3 microns, the mean free path for these electrons will be considerably greater than for those in the crystals, and consequently, their contribution to the conductivity per electron is greatly increased. The question as to whether the conductivity of the pores will contribute a significant portion to the total conductivity of the aggregate depends on the relative magnitudes of the above-mentioned parameters.

Loosjes and Vink calculated the conductivity of the pores by first computing the density of electrons in the pores, and then considering the increase in the mean velocity of the electrons under the influence of an electric field acting for the mean period of time required for the electrons to travel their mean free path. The results were not expressed by a single compact equation, but required a separate series of calculations for each individual case. In what follows, the conductivity of the pores will be calculated in a somewhat more direct manner that does not necessitate the extensive use of mean values. The results are in qualitative agreement with those of Loosjes and Vink and have the advantage of being expressed in a more easily applied form.

In arriving at a basis for calculating the conductivity of the pores, the structure is visualized as an irregular aggregate of crystals all in contact with each other and with the interstices between the crystals interconnecting. The mean density of the aggregate for oxide cathodes is about 25 to 50 percent of crystal density. When an electric field is applied across the aggregate, the electrons thermionically emitted from one side of a pore will be in an accelerating field while those emitted from the opposite side of the pore will be in a retarding field, the result being a net current passing across the pore. In what follows, two types of approximation to the geometry of a pore will be made, each being an extreme case of the real situation.

II. INFINITE PARALLEL PLANES MODEL FOR AN INTERSTICE

In this first approximation to the geometry of a conducting pore, it is assumed that the surfaces normal to the electric field represent the essential sources of the current and that the side walls of the pore may be omitted from the calculation. This situation may be represented by two flat surfaces of infinite extent possessing equal thermionic work functions, equal temperatures and spaced s meters apart. The energy level diagram for this case is shown in Fig. 1. Both surfaces have a work function $e\phi$. eV_m is the maximum potential energy of an electron resulting from the space charge and is measured relative to the top of the field free electron affinity of the left-hand surface. The net current density crossing eV_m will be equal to the current density crossing from the left minus the current density

* Supported in part by the ONR.
† Now at Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts.
[1] For review articles see J. P. Blewett, J. Appl. Phys., **10**, 668–679, 830–848 (1939); **17**, 643–647 (1947). A. S. Eisenstein, *Advances in Electronics* (Academic Press, Inc., New York, 1948), Vol. 1, Part 1.
[2] R. Loosjes and H. J. Vink, Philips Research Rept., **4**, 449–475 (1949).

crossing from the right. Thus,

$$J = J_0(1-\epsilon^{-Ees/kT})\epsilon^{-eV_m/kT}, \quad (1)$$

where J_0, the zero field thermionic emission, is given by the Richardson-Dushman equation:

$$J_0 = \frac{4\pi mek^2T^2}{h^3}\epsilon^{-e\phi/kT}. \quad (2)$$

For small fields E, such that $Ees < kT$, the expression inclosed in the parenthesis may be approximated by series expansion, only the first power terms being retained. The conductivity of the pore σ may then be obtained by dividing the current density by the electric field:

$$\sigma = \frac{J_0 es}{kT}\epsilon^{-eV_m/kT}. \quad (3)$$

The extent to which the space charge potential eV_m influences the conductivity of this model may be estimated with the aid of a transcendental equation derived by Loosjes and Vink[2] for the case of zero applied electric field[3]:

$$\frac{es}{4kT}\left[\frac{2J_0}{e\epsilon_0}(2\pi mkT)^{\frac{1}{2}}\right]^{\frac{1}{2}} = \epsilon^{eV_m/2kT}\arctan(\epsilon^{eV_m/kT}-1)^{\frac{1}{2}}. \quad (4)$$

Since all the quantities on the left of Eq. (4) are assumed to be known, we may find eV_m/kT by graphing the right side of Eq. (4) as a function of eV_m/kT. Typical values for V_m are given in Table I.

In the derivation of Eq. (3) we have neglected to take into account the Schottky effect. This effect of the applied field lowering the barrier of the left-hand surface will be significant when the space charge hump is negligible. In the absence of a space charge, Eq. (1) predicts that the current-voltage characteristics will be concave toward the voltage axis and will tend to saturate at a current density equal to the zero field thermionic emission. However, due to the Schottky effect the current will not saturate but will continue to rise with a slope which should be in qualitative agreement with the Schottky theory. In fact, we should expect this slope to be somewhat greater than that predicted by the Schottky theory because the actual surfaces involved in the pores of the cathode aggregate are not flat but are rough and nonuniform.

TABLE I. Typical values of the space charge maximum V_m in volts as derived from Eq. (4).

J_0 in amp/cm²	$s=1\mu$ $T=700°K$	$s=1\mu$ $T=1000°K$	$s=5\mu$ $T=1000°K$
0.01	0.000035	0.000030	0.00069
0.1	0.00034	0.0029	0.0066
1.0	0.0032	0.0028	0.046
10.0	0.025	0.024	0.155

[3] Note: ϵ_0 is the permittivity of vacuum; all the other symbols have their usual meanings. The units used are rationalized mks except where specifically indicated otherwise.

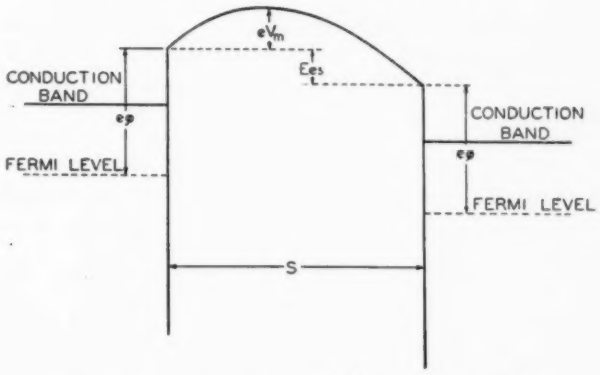


FIG. 1. Energy level diagram for the infinite parallel planes model of an interstice.

A further effect to be mentioned is with regard to the behavior of the space charge in the presence of an applied field. An attempt was made to derive a simple relationship for V_m in the presence of an applied field by a straightforward extension of the method employed by Loosjes and Vink in obtaining Eq. (4). However, the result could only be evaluated by extensive numerical integration and then only after several approximations had been made. For small values of the field E, the reduction in V_m can be approximated by simply subtracting $Es/2$, the value of the superimposed potential at a point midway between the surfaces. For higher fields the reduction in V_m is less rapid until V_m is reduced to zero, after which the surface barrier will begin to be lowered by a modified Schottky effect. The net result is that the curvature of the current-voltage characteristic is reduced by the presence of the space charge so that at higher temperatures the approximately linear region will be greatly extended. These conclusions regarding the shape of the current-voltage characteristics are in qualitative agreement with the results obtained by Loosjes and Vink.[2]

III. INFINITE RECTANGULAR PRISM MODEL FOR AN INTERSTICE

The principal criticism of the parallel planes model for calculating the conductivity of a pore is that it neglects the effect of the side walls. It would not be unreasonable to suspect that the electrons emitted from these side walls could produce a substantial increase in the magnitude of the conductivity, especially in a long thin pore in which the electric field lies along the long dimension. In order to investigate this possibility a second model will be considered, namely, a rectangular prism in which the z dimension extends to infinity in both directions. Figure 2 shows a cross section of this model. A uniform electric field is assumed to lie along the y axis so that the force on an electron will be in the positive direction. The width (x dimen-

sion) is W and the height is s. The z dimension has been taken as infinite in order to simplify the calculations.

The net current per unit z crossing a surface parallel to the $x-z$ plane and y meters above the bottom of the cavity is given by the following equation:

$$i=JW=\int_{-W/2}^{W/2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \times eV_y f(x, y, z, V_x, V_y, V_z)dV_x dV_y dV_z dx, \quad (5)$$

where J is the mean current density and $f(x, y, z, V_x, V_y, V_z)$ is the density of electrons in (velocity) phase space. The assumed boundary conditions are that the electrons are emitted with a Maxwellian distribution of velocities. That is, at the pore walls

$$f_0=D\exp\left\{-\frac{m}{2kT}[V_{x0}^2+V_{y0}^2+V_{z0}^2]\right\}, \quad (6)$$

where

$$D=\frac{2m^3}{h^3}e^{-e\phi/kT}.$$

On leaving the walls the electrons experience a constant acceleration of α meters/sec^2. Analysis[4] of this situation shows that f must be expressed in four separate equations corresponding to the four different surfaces from which the electrons are emitted. Using the numbers assigned to the four surfaces in Fig. 2, it may be shown that the density of electrons for the corresponding four regions in phase space is

$$f_1=D\exp\left\{-\frac{m}{2kT}[V_x^2+V_y^2+V_z^2-2\alpha y]\right\},$$

$$f_2=D\exp\left\{-\frac{m}{2kT}[V_x^2+V_y^2+V_z^2-2\alpha(y-s)]\right\},$$

$$f_3=D\exp\left\{-\frac{m}{2kT}\left[V_x^2+V_y^2+V_z^2+\alpha^2\left(\frac{x-W/2}{V_x}\right)^2-2\alpha V_y\left(\frac{x-W/2}{V_x}\right)\right]\right\},$$

$$f_4=D\exp\left\{-\frac{m}{2kT}\left[V_x^2+V_y^2+V_z^2+\alpha^2\left(\frac{x+W/2}{V_x}\right)^2-2\alpha V_y\left(\frac{x+W/2}{V_x}\right)\right]\right\}. \quad (7)$$

In order to substitute Eq. (7) into Eq. (5) it is necessary to know the limits of the regions in phase space to which they apply. The boundaries of these regions will be functions of the four variables x, y, V_x, V_y. The x and y coordinates will range over the entire cross section of the cavity for each of the four functions; hence, we need only state the regions in the V_xV_y-plane for each point x, y. Using the fact that the V_x, V_y velocities of an electron at a point x, y uniquely determine the surface from which it originated, it may be shown[3] that the boundaries for region II are

$$-\infty<V_y<(2\alpha y)^{\frac{1}{2}}$$

$$V_a\equiv\frac{[x-(W/2)]\alpha}{V_y+[V_y^2+2\alpha(s-y)]^{\frac{1}{2}}}<V_x<\frac{[x+(W/2)]\alpha}{V_y+[V_y^2+2\alpha(s-y)]^{\frac{1}{2}}}\equiv V_b. \quad (8)$$

And similarly the boundaries for region I are

$$(2\alpha y)^{\frac{1}{2}}<V_y<\infty$$

$$V_c\equiv\frac{[x-(W/2)]\alpha}{V_y-(V_y^2-2\alpha y)^{\frac{1}{2}}}<V_x<\frac{[x+(W/2)]\alpha}{V_y-(V_y^2-2\alpha y)^{\frac{1}{2}}}\equiv V_d. \quad (9)$$

Regions III and IV, in which f_3 and f_4 apply, constitute the remainder of the V_xV_y-plane with V_x, respectively, negative and positive. Equation (5) may now be written

$$JW=e\int_{-W/2}^{W/2}\int_{-\infty}^{\infty}\left\{\int_{-\infty}^{(2\alpha y)^{\frac{1}{2}}}V_y\left[\int_{-\infty}^{V_a}f_3dV_x+\int_{V_a}^{V_b}f_2dV_x+\int_{V_b}^{\infty}f_4dV_x\right]dV_y+\int_{(2\alpha y)^{\frac{1}{2}}}^{\infty}V_y\left[\int_{-\infty}^{V_c}f_3dV_x+\int_{V_c}^{V_d}f_1dV_x+\int_{V_d}^{\infty}f_4dV_x\right]dV_y\right\}dV_x dx. \quad (10)$$

Carrying out the indicated integrations in Eq. (10) is a rather long procedure. However, by assuming that

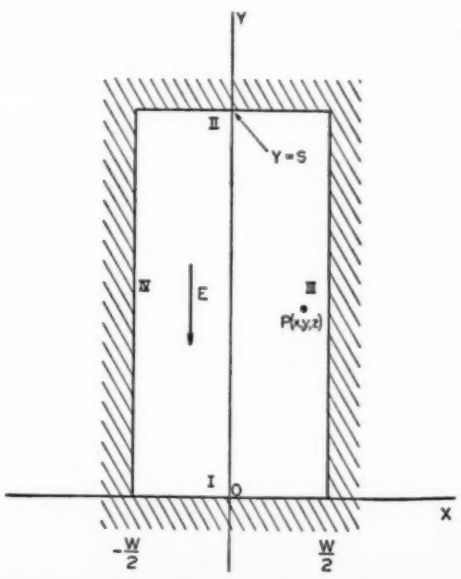


FIG. 2. Cross section of rectangular prism model of an interstice.

[4] Eugene B. Hensley, Ph.D. thesis, University of Missouri.

the electric field is small such that $Ees<kT$ and also assuming that $W\ll s$, the following approximate solution has been obtained for the case where $y=s/2$:

$$J=\frac{J_0EeW}{kT}\left(\frac{1}{8}+\frac{1}{4}\ln\frac{s}{W}\right). \tag{11}$$

Equation (11) is to be compared with Eq. (3). Since $W\ll s$, the model used in deriving Eq. (11) is essentially composed of two parallel surfaces W meters apart. However, this model differs from the one in Sec. II in that the electric field is parallel to these surfaces instead of being perpendicular to them. Thus, these two models represent two extreme cases for conducting pores. However, it should be noted that in both cases the specific conductivity is controlled primarily by the smallest dimension of the pore. Thus Eq. (11) has the same general form as Eq. (3) [omitting the space charge factor in (3)] except for the slowly varying quantity in the bracket of Eq. (11). Since this quantity is usually less than one, it may be concluded that the net effect of the side walls is to reduce the current density rather than to increase it. These results may be summarized by stating that the order of magnitude of the specific conductivity of the pores may be expressed by

$$\sigma_p=J_0el_p/kT, \tag{12}$$

where l_p is some sort of average diameter of the pores weighted in favor of the narrower dimensions. The effect of the space charge at the higher temperatures will be to reduce this conductivity in a manner somewhat similar to that indicated in Sec. II. In fact, the form indicated in Eq. (3) may be expected to represent the maximum reduction, with the actual case being somewhat less.

IV. APPARENT CONDUCTIVITY OF POROUS SEMICONDUCTORS

The manner in which the specific conductivities of the pores and crystals combine to form the apparent conductivity of the aggregate can only be treated in a qualitative sort of way. Since the electron affinity of the crystals is probably always positive, the temperature dependence of the pore conductivity will be greater than that of the crystals. Consequently, we may expect the conduction to be through the pores at high temperatures and through the crystals at low temperatures.

We will first consider an extreme case in which the temperature is sufficiently high that the conductivity of the crystals is negligible. The pores are assumed to be interconnected so that the conduction current may proceed through the aggregate by being successively absorbed and reemitted by the pore walls. The net current at a given surface will be negligible since it will be receiving as much current from earlier surfaces as it re-emits to later surfaces. This is a self-adjusting situation, since if it were not true for a given surface, an electrostatic charge would accumulate and alter the electric field in such a manner as to make it true. Thus we may visualize the flow lines of the current density as threading their way through the porous aggregate from one electrode to the other. Also, the electric field will have become distorted so that it is everywhere more or less parallel to the lines of J. From this it may be seen that the model discussed in Sec. III more closely represents this case than does the parallel plane model discussed in Sec. II. However, since both models lead to similar formulas, Eq. (12) should be a good approximation. While Eq. (12) represents the conductivity of the pores, the apparent conductivity of the aggregate will be smaller than this by an unknown geometric factor which we may represent by the symbol $\alpha_p<1$.

In the opposite extreme will be the case of low temperatures in which the pore conductivity is negligible compared with that of the crystals. Again the electric field will be distorted in such a manner that it follows along approximately parallel to the lines of the conduction current density. The apparent conductivity in this case will be lower than the specific conductivity of the crystals by an unknown factor $\alpha_c<1$ which will depend on the geometry of the crystals and on the degree of contact between the crystals.

At temperatures intermediate to the two extreme cases mentioned above, a portion of the conduction current will pass in series through the pores and crystals. In order to clarify this situation an artificial case will be considered in more detail. We will assume that the apparent conductivity of an aggregate may be expressed in three terms. The first term will be for the current conducted only by the pores, the second for the current passing only through the crystals, and the third term for the series current. We will also assume that all the lines of the series current density pass through the pores R times as much as through the crystals. The apparent conductivity for this model may be written

$$\sigma=\alpha_p\sigma_p+\alpha_c\sigma_c+\alpha_s(R+1)\frac{\sigma_p\sigma_c}{\sigma_p+R\sigma_c}, \tag{13}$$

where α_s is the geometric factor for the series term.

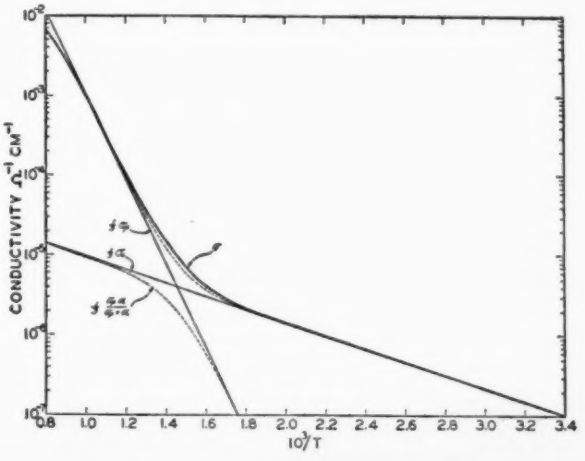


FIG. 3. Conductivity of a porous semiconductor.

The conductivity plot in Fig. 3 was made using Eq. (13) together with the following typical values for the parameters: σ_p was expressed by Eq. (12) taking $J_0=1$ amp/cm^2 at 1000°K and with an apparent work function $e\phi_0=1$ ev. The mean pore size l_p was taken as 3 microns. The conductivity of the crystals σ_c was expressed by the equation

$$\sigma=\sigma_0 e^{-Q/kT}, \tag{14}$$

in which $\sigma_0=0.013\Omega^{-1}m^{-1}$ and $Q=0.2$ ev.[5] The α's were chosen arbitrarily to be $\alpha_p=\alpha_c=2\alpha_s=\frac{1}{3}$ and $R=1$. The effect of the space charge was estimated by using the space charge factor in Eq. (3) together with values of V_∞ obtained from Eq. (4). However, this cannot be considered as more than an indication that the conductivity will be somewhat reduced at the higher temperatures.

Two important features are to be noted regarding Fig. 3. First, if the curve of σ had been represented by a series of points and if an attempt had been made to pass two intersecting straight lines through these points, the resulting temperature dependence of the conductivity in the pore conduction region would have been interpreted as being slightly less than that of σ_p. Second, the dotted curve just below σ represents the conductivity as it would have been if the series term had been omitted. Thus it is seen that the series term makes only a small contribution to the general shape of the curve and this principally in the region of the break.

We may now discuss Eq. (13) in relation to a real porous semiconductor. α_p and α_c will very probably be temperature dependent. α_p will have its maximum value at high temperatures and will decrease at lower temperatures as the distortion of the electric field becomes less favorable for continuous conduction through the pores. σ_c may have its maximum value at low temperatures; however, the effect of thermal expansion on the contacts between the crystals may be the determining factor. With regard to the series conduction term, the assumption of a single value for R was of course extremely unrealistic. However, if a long series of such terms had been used, the individual geometric factors α_s would, of necessity, have been very small. Since terms of this type have their maximum influence in the vicinity of the break in the conductivity curve, as is indicated by the results in Fig. 3, it may be supposed that the net result is a compensating effect on the decrease in the values of α_p and α_c in this region. Consequently, assuming α_p and α_c constant (this is essentially what Loosjes and Vink have done) probably will not lead to serious error. However, very little can be predicted regarding the exact shape of the curve in the middle temperature region, and in particular the slopes should be regarded as very qualitative.

[5] Note: In terms of equation (16) and (20) this is equivalent to $\mu_0=-0.2$ ev, $\mu'=-10^{-3}$ ev/°K and $l_c=5$ angstroms.

V. RATIO OF CONDUCTIVITY TO THERMIONIC EMISSION

The ratio of conductivity to thermionic emission for pore conduction may be obtained from Eq. (12):

$$\sigma_p/J_0=el_p/kT. \tag{15}$$

It is interesting to note that this ratio is a function of the temperature and geometry of the porous aggregate, but to the extent that the conduction is through the pores, it is independent of the composition of the crystals. For a mean pore size of 3 microns and a temperature of 1000°K this ratio is 3.5×10^{-3} cm/ampΩ. This, however, does not take into account the correction factors relating the apparent conductivity and thermionic emission to their specific values.

The corresponding ratio for crystal conduction will depend on the assumptions made in deriving an equation for the conductivity. If the usual assumption is made, that the mean free path of the electrons in the conduction band of the semiconductor is independent of their energy, the equation for the conductivity is

$$\sigma_c=\frac{16\pi mkTe^2l_c}{3h^3}e^{\mu/kT}, \tag{16}$$

where μ is the Fermi energy measured from the bottom of the conduction band, which is designated as the zero energy. Since we are dealing with nondegenerate semiconductors, this means that μ is always a negative number. Dividing Eq. (16) by the Richardson-Dushman Eq. (2) we obtain for crystal conduction

$$\frac{\sigma_c}{J_0}=\frac{4}{3}\frac{el_c}{kT}e^{\chi/kT}, \tag{17}$$

where χ is the electron affinity of the semiconductor and where

$$e\phi=\chi-\mu. \tag{18}$$

The mean free path l_c is believed to be of the order of the interatomic distances for polar crystals.[6] Hence Eq. (17) tells us that a key question regarding the possibility of pore conduction in porous semiconductors, such as the alkaline earth oxides, is the magnitude of the electron affinity χ. If $\sigma_c/J_0<\sigma_p/J_0$, then pore conduction will predominate and vice versa. For example, if l_c is 5 angstroms, the value of χ which will result in $\sigma_c/J_0=\sigma_p/J_0$ is 0.7 ev at 1000°K. If l_c is 50 angstroms, χ would only need to be 0.5 ev. Of course, the geometrical factors including the reflection coefficient for the emission equation must be considered in comparing experimental values of σ/J_0 with Eqs. (15) and (17). Also it is necessary that the measurements of σ and J be made on a single sample which is in a given state of activation.

Few reliable values for σ/J appear in the literature.

[6] N. E. Mott and R. W. Gurney, *Electronic Processes in Ionic Crystals* (Oxford University Press, Oxford, England, 1940).

Hannay, MacNair, and White[7] have reported rather extensive data taken at a fixed temperature of 970°K giving a value of 0.025 cm/ampΩ. This is seen to be about an order of magnitude larger than is predicted by the pore conduction hypothesis, and it would appear questionable as to whether the unknown geometrical factors could compensate for the difference. However, the following feature may be of great importance in measurements of this type. It is known that metallic barium is evaporated from the outer surfaces of oxide cathodes. Since the activity of these cathodes is believed to depend on a stoichiometric excess of barium, this may result in a reduction in the activation of the outer surface with a consequent reduction in the measured value of the thermionic emission. Barium evaporated from the walls of the pores would not be lost so that this effect would not be present in the interior of the cathode coating. Further experimentation on this point would seem desirable.

It would be very helpful in deciding the existence of pore conduction if the magnitude of the electron affinity χ could be determined, especially since this quantity should be independent of the activation of the cathode, assuming that dipole layers on the surface are not of fundamental importance. However, it should be noted that χ cannot be obtained by comparing the slopes of a Richardson plot with that of a conductivity plot even though pore conduction is absent. This is a misconception that has occasionally appeared in the literature. The thermionic work function of a semiconductor must be regarded as a temperature dependent quantity which can, however, be expressed by the first two terms of a power series over reasonable ranges of temperatures:

$$e\phi = e\phi_0 + e\phi' T. \tag{19}$$

Substituting Eq. (19) into Eq. (2) shows that the slope of a Richardson plot is related to $e\phi_0$ and not to $e\phi$. Similarly the Fermi energy and the electron affinity may be represented as linear functions of temperature over limited temperature ranges

$$\mu = \mu_0 + \mu' T, \tag{20}$$

$$\chi = \chi_0 + \chi' T. \tag{21}$$

Hutner, Rittner, and DuPre[8] have given several numerical examples illustrating the manner in which μ behaves as a function of temperature for various distributions of acceptor and donor centers in semiconductors. Substituting Eq. (20) into Eq. (16) shows that the slope of a conductivity plot for a semiconductor is related to μ_0 and not to μ. Thus since only μ_0 and $e\phi_0$ are known,[9] Eq. (18) cannot be used. However, the equation

$$e\phi_0 = \chi_0 - \mu_0 \tag{22}$$

will give us a value for χ_0. Thus, for example, the experimental data for oxide-coated cathodes usually show $e\phi_0$ and μ_0 (using the semiconductor hypothesis) to be very nearly equal at high temperatures so that χ_0 is usually interpreted as being of the order of one- or two-tenths of an electron volt. It must be emphasized that this is a measure of χ_0 and not χ, and it is this only if the conduction at high temperatures is through the crystals. However, as was pointed out near the end of Sec. IV, the slight difference in the high temperature slopes is also to be expected if the pore conduction hypothesis is correct.

A further observation can be made which lends support to the pore conduction hypothesis for oxide cathodes. As was indicated in Fig. 3, conductivity plots for well-activated oxide cathodes show two regions of linearity. Equations (16) and (20) can be fitted to each of these regions separately providing μ_0 in Eq. (20) is selected to correspond with the particular slope. However, this necessity of using two values of μ_0 (and consequently two values of μ'), together with Eq. (18), implies that the Richardson plot will have a corresponding change in slope. Experimentally, however, Richardson plots are always straight in this region. Perhaps the most striking example has been obtained by Sparks[10] from a cathode consisting of pure BaO on pure nickel. The Richardson plot for this cathode produced a straight line for more than 12 orders of magnitude of current and over a temperature range from 368°K to 1017°K. The apparent work function $e\phi_0$ was 1.37 ev and the A constant was 0.122 amp/cm^2 °K^2. From this type of data we may conclude that both slopes of the conductivity plots cannot be explained as arising from electronic conduction through semiconducing crystals.

VI. THERMOELECTRIC EFFECT IN POROUS SEMICONDUCTORS

There are two sources of thermal emf in a porous semiconductor; one is in the crystals of the aggregate and the other is in the interstices between the crystals. The theory of the thermoelectric power of semiconductors is well-known. However, most of the formulas to be found in the literature are based on definite assumptions regarding the type of impurity centers involved. In keeping with the form of the equations in Sec. V, it is desirable to express the thermoelectric power in terms of a temperature dependent Fermi energy [Eq. (20)], thus avoiding the necessity of any assumptions regarding the impurity centers.

The thermoelectric power θ' is related to the Peltier coefficient $\pi_{1\to 2}$ by the relation

$$\theta' = \pi_{1\to 2}/T. \tag{23}$$

A general expression for the Peltier coefficient has been

[7] Hannay, MacNair, and White, J. Appl. Phys. **20**, 670 (1949).
[8] Hutner, Rittner, and DuPre, Philips Research Rept. **5**, 188–204 (1950).
[9] Note: We may also determine $e\phi'$ by making assumptions regarding the reflection coefficient in the Richardson-Dushman equation.
[10] I. L. Sparks, Ph.D. thesis, University of Missouri, 1951.

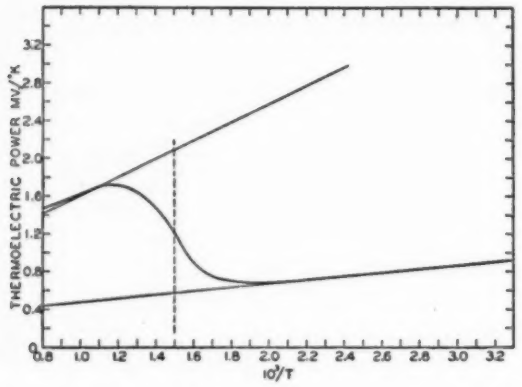


FIG. 4. Thermoelectric power of a porous semiconductor.

given by Seitz:[11]

$$\pi_{1\to 2}=\frac{T}{e}\left[\left(\frac{K_2}{TK_1}-\frac{\mu}{T}\right)_2-\left(\frac{K_2}{TK_1}-\frac{\mu}{T}\right)_1\right], \quad (24)$$

where

$$K_i=\frac{C}{V}\int_0^\infty \epsilon^i l\frac{\partial f(\epsilon)}{\partial \epsilon}d\epsilon, \quad (25)$$

$f(\epsilon)$ is the Fermi distribution function, and l is the mean free path of the electrons in the conductor. As applied here, the subscripts 1 and 2 on the parentheses refer to the metal and semiconductor of the junction, respectively. Since the thermoelectric power of a metal is of the order of 10^{-3} times that of a semiconductor, the second parenthesis may be neglected. In order to evaluate Eq. (25) for the semiconductor term, it is usually assumed that the mean free path is independent of the energy. Using the classical approximation for $f(\epsilon)$ it may then be shown that

$$K_2/TK_1=2k. \quad (26)$$

The thermoelectric power for an n-type semiconductor is thus found to be

$$\theta_c'=-\frac{1}{e}\left(2k-\frac{\mu}{T}\right). \quad (27)$$

If Eq. (20) is used for μ, this becomes

$$\theta_c'=-\frac{1}{e}\left(2k-\mu'-\frac{\mu_0}{T}\right). \quad (28)$$

Plotting θ_c' as a function of $1/T$ should result in a straight line whose slope is dependent on μ_0. It will be recalled from Sec. V that μ_0 also determines the slope of the conductivity plot.

The thermoelectric power of a pore will be derived on the basis of the parallel planes model of an interstice discussed in Sec. II. Referring to Fig. 1, if the right-hand surface is at a temperature T and the left-hand surface at $T+\Delta T$, then a potential $Es=\Delta\theta$ will be required to reduce the net current between the surfaces to zero. Equating the emission current densities from the two surfaces, we have

$$AT^2e^{-e(\phi+V_m)/kT}$$
$$=A(T+\Delta T)^2e^{-e(\phi\to\phi'\Delta T+V_m+\Delta\theta)/k(T+\Delta T)}. \quad (29)$$

Taking the logarithm of both sides, substituting Eq. (19), and rearranging, we obtain

$$\frac{2k}{e}\ln\frac{T+\Delta T}{T}=\frac{\phi_0+V_m+\Delta\theta}{T+\Delta T}-\frac{\phi_0+V_m}{T}, \quad (30)$$

which may be solved for $\Delta\theta/\Delta T$:

$$\frac{\Delta\theta}{\Delta T}=\frac{2k}{e}\frac{(T+\Delta T)}{\Delta T}\ln\left(\frac{T+\Delta T}{T}\right)+\frac{\phi_0+V_m}{T}. \quad (31)$$

In the limit at $\Delta T\to 0$, we find the thermoelectric power to be

$$\theta_p'=\frac{d\theta}{dT}=\frac{1}{e}\left[2k+\frac{e(\phi_0+V_m)}{T}\right]. \quad (32)$$

Remembering that μ is a negative quantity, Eq. (32) is seen to resemble Eq. (27). The space charge potential V_m may be estimated with the aid of Eq. (4). When θ_p' is plotted as a function of $1/T$, a straight line should result whose slope is dependent on $e\phi_0$, the apparent work function of the semiconductor. Of course, the space charge may be expected to reduce this slope at high temperatures.

The two sources of thermoelectric power discussed above are essentially in parallel with each other. This situation may be regarded as one in which each pore is in parallel with its adjacent crystals, or, alternatively, as an entire system of interconnecting pores in parallel with the crystal matrix; the result will be essentially the same in either case. At extreme temperatures, the mechanism which has the lowest internal resistance will control the effective thermoelectric power θ' of the aggregate, but at intermediate temperatures an average between the two emf's will result.

This situation may be analyzed by considering θ_p' as a voltage generator in series with a conductance σ_p, and in parallel with this a voltage generator θ_c' in series with a conductance σ_c. θ', the effective thermoelectric power of the aggregate, will be the resulting voltage across this parallel system. Solving for θ' we find

$$\theta'=\frac{\sigma_p}{\sigma_p+\sigma_c}\theta_p'+\frac{\sigma_c}{\sigma_p+\sigma_c}\theta_c'. \quad (33)$$

In this equation, $\sigma_p+\sigma_c$ will be proportional to σ, the effective conductivity of the aggregate. Figure 4 has been drawn from Eq. (33), using the same parameters

[11] F. Steitz, *The Modern Theory of Solids* (McGraw-Hill Book Company, Inc., New York, 1940), p. 180.

that were used in producing Fig. 3. The upper straight line represents the temperature dependence of θ_p' and the lower straight line represents θ_c'. The dotted vertical line indicates the temperature at which the break in the conductivity occurs.

The experimental data on the thermoelectric power of oxide cathodes by Young[12] are in good qualitative agreement with this analysis, with the possible exception of the slope in the low temperature region. One possible explanation of why this slope is not as large as would be expected is that there is a third conduction mechanism in parallel with the cathode. This could either be leakage between the electrodes of the tube or possibly a third mechanism in the cathode itself, such as ionic conduction. Analysis of this situation shows that such an additional path for the electric current would produce just such a decrease in this low temperature slope and could even reverse it.

The author wishes to express his gratitude to Professor G. H. Vineyard under whose guidance this work was undertaken and completed, and to Professor A. S. Eisenstein for helpful suggestions.

[12] J. R. Young, Ph.D. thesis, University of Missouri, 1951, and J. Appl. Phys. **23**, 1129 (1952).

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 10 OCTOBER, 1952

Electrical Conductivity and Thermoelectric Power of (BaSr)O and BaO*

J. R. YOUNG†
Department of Physics, University of Missouri, Columbia, Missouri
(Received March 31, 1952)

A study has been made of the thermoelectric power, the electrical conductivity, and the thermionic emission of (BaSr)O and BaO at different states of activation and on different base metals over a temperature range of 1100°K to 300°K. A plot of the logarithm of the conductivity *vs* reciprocal temperature showed two straight line regions. The temperature dependence of the thermionic emission was found to be the same as the temperature dependence of the conductivity in the temperature range 700°K to 1100°K. The slope of a plot of the thermoelectric power *vs* reciprocal temperature was also similar to the thermionic work function and the conductivity activation energy in this temperature range.

The thermoelectric power varied with temperature from a value of about 2.0 mv/° at 1100°K, increasing to about 2.5 mv/° at 800°K, reducing to between 1.0 and 0.5 mv/° at about 500°K, and remaining practically constant or increasing slowly at still lower temperatures. In all thermoelectric power measurements the hot junction was positive indicating negative charge carriers.

The simple semiconductor theory does not adequately account for these results which seem to be in good agreement with the pore conduction hypothesis.

I. INTRODUCTION

THE impurity semiconductor model, first proposed by Wilson[1] in 1931, has been successful in accounting for many properties of the alkaline earth oxides, although the simple one impurity level model does not satisfactorily explain all of the experimental observations. Recently a new theory was proposed by Loosjes and Vink[2] to explain the conductivity of oxide cathodes. They consider two parallel conduction mechanisms with quite different temperature dependence, each predominating in different temperature regions. At low temperatures, below 800°K, the conduction is thought to be primarily electronic through the crystal grains; at high temperatures, above 800°K, the current is carried by an electron gas in the pores of the cathode. The density of electrons in these pores will depend upon the electron emission from the oxide; therefore the temperature dependence of the conductivity should be the same as that of thermionic emission. Hensley[3] has recently considered the properties of a porous semiconductor and has arrived at theoretical expressions predicting the magnitude of the conductivity and thermoelectric power and their temperature dependence.

When probes are embedded in oxide cathode coatings for aid in measuring the electrical conductivity, it is frequently found that a potential difference exists between the probe and the base metal when no current is drawn. This was attributed by Becker and Sears[4] to a thermoelectric voltage arising from the temperature gradient in the coating. Except for this and similar observations no systematic study of the thermoelectric effect in the alkaline earth oxides is reported in the literature.[5] Such a study with simultaneous measurements of conductivity and thermionic emission should provide evidence for or against the predictions of the above-mentioned theories. It was the purpose of this

* Supported in part by the ONR.

† Now at the General Electric Research Laboratory, Schenectady, New York.

[1] A. H. Wilson, Proc. Roy. Soc. (London) **A134**, 277–287 (1931).

[2] R. Loosjes and H. J. Vink, Philips Research Rept. **4**, 449–475 (1949).

[3] E. B. Hensley, thesis, University of Missouri (1951); J. Appl. Phys. **23**, 1122 (1952).

[4] J. A. Becker and R. W. Sears, Phys. Rev. **38**, 2193–2213 (1931).

[5] A recent study published after this report was in preparation, Kin-ichi Noga, J. Phys. Soc. Japan **6**, 124–125 (1951).

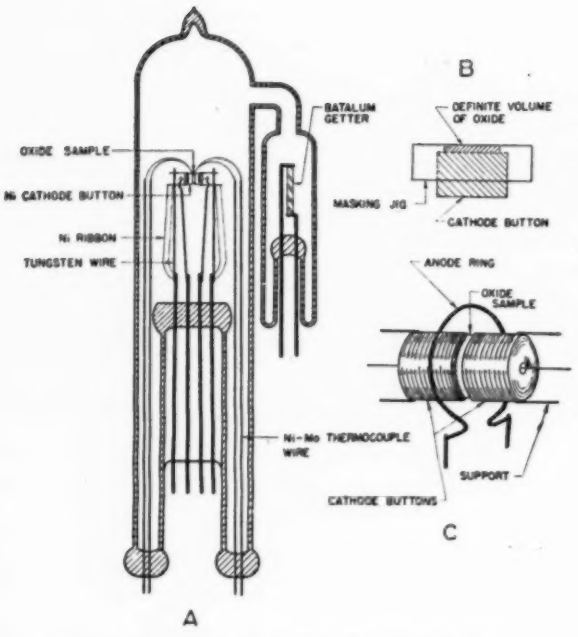


FIG. 1. Schematic diagram of experimental tube.

investigation to study the temperature dependence of the electrical conductivity, thermoelectric effect, and thermionic emission of (BaSr)O and BaO at different stages of activation and on different base metals.

II. EXPERIMENTAL PROCEDURE

A. Tube Design and Construction

The most suitable electrode geometry for the simultaneous measurements of conductivity and thermoelectric power was thought to be that of Loosjes and Vink[2] upon which the tube design shown in Fig. 1 is based. For the simultaneous measurements of conductivity, thermoelectric power, and thermionic emission a ring anode was introduced (Fig. 1C). The flat cathode buttons‡ were 3.6 mm in diameter of the type used in the 2C36 tube. The nickel stock used in these buttons will be referred to as standard nickel. In some of the tubes 1001 pure nickel§ disks 0.006 in. thick were welded over the standard nickel cathode buttons. Impurity content of the two types of nickel used as base metals was determined by spectrographic analysis and is given in Table I.

The samples were prepared by spraying two previously processed (see Sec. II B) cathode buttons. The coatings used were equal molar $(BaSr)CO_3$|| or pure $BaCo_3$.¶ The impurity content of the coatings used, as determined from spectrographic analysis, is shown in Table I. In order to obtain a coating of uniform thickness and area, a stainless steel masking jig was machined to fit over the cathode button before spraying (Fig. 1B). The carbonate was sprayed over the jig to a thickness greater than ultimately desired; the excess was then scraped off flush with the top of the jig with a razor blade. This resulted in each of the cathode buttons having a centered coating 0.15-mm thick and 4.4×10^{-2} cm² in area. Each cathode button was supported by two 0.030-in. nickel wires. These wires were welded to a 0.025-in. tungsten wire which served as a spring to keep the two sprayed buttons pressed tightly together. A nickel ribbon 0.003 in. thick and approximately 3 mm wide was used as a jumper around the tungsten spring. This prevented the tungsten spring from heating and losing its spring-like properties due to the passage of heater current.

The temperature of each nickel button was measured by a thermocouple made of 0.002-in. molybdenum and nickel wires welded, before spraying, to the edge of each button. These wires were then welded to 0.040-in. molybdenum and nickel support leads which were attached to the two lead extended presses at the bottom of the tube just above the glass to metal seals. This cold junction of the thermocouple was considered to be at room temperature. The Mo—Ni thermocouple wire was calibrated against a Pt—Pt Rh thermocouple calibrated at the Bureau of Standards to an accuracy of 0.1°C.

B. Tube Processing

Before spraying, the cathode buttons were chemically cleaned and hydrogen fired at 900°C for fifteen minutes as were all nickel support leads and the 1001 nickel disks used for facing the buttons in some tubes. After hydrogen firing, the heaters were installed and the cathode buttons were assembled and mounted in a separate glass envelope for vacuum firing at 1000°C until the pressure, as indicated by an ionization gauge, was approximately 10^{-8} mm of mercury. The cathode buttons were then sprayed and the final tube assembled as described in the previous section.

The tubes were evacuated by a three-stage Distillation Products oil diffusion pump and were baked at 450°C for one hour. The ultimate vacuum was approximately 10^{-8} mm of Hg. The cathodes were heated slowly to 1150°K. They remained at this temperature until the carbonate was completely converted to the oxide and the pressure had decreased again to about 10^{-8} mm of Hg.

Activation was accomplished by passing a direct current between the buttons, through the coating, first in one direction and then in the other. Less than 6 volts was used in this process. Considerable polarization was observed in the early stages of activation. However, when high activity was reached no polarization was apparent. Some of the tubes were completely

‡ Obtained from Sylvania Electric Products, Inc., Kew Gardens, Long Island, New York.

§ 1001 nickel was obtained from E. M. Wise, International Nickel Company, New York, New York.

|| C51–2 obtained from Ratheon Manufacturing Company, Newton, Massachusetts.

¶ Ultra-Pure $BaCO_3$ from Mallinckrodt Chemical Works, St. Louis, Missouri.

TABLE I. Impurity content in parts per million of the coatings and base metals used.

	Fe	Mn	Pb	Cu	Na	Ca	Sr	Mg	Si	Co	Al	Ag
$BaCO_3$	3	0.5	0.5	0.5	10	5	30	50	[a]	...	[a]	...
$(BaSr)CO_3$	...	20	10	2	...	...	...	500	2	...	10	...
1001 Ni	10	1	10	10	...	...	...	2	...	100	...	[a]
Standard Ni	20	50	5	50	...	...	...	50	10	100	...	[a]

[a] Possible trace.

activated on the vacuum system while others were activated after seal-off. All tubes were sealed off the vacuum system at a pressure of approximately 10^{-8} mm of Hg following which the Batalum getter was flashed.

C. Method of Measurements

Dc voltage-current curves were found to be linear at all temperatures and at all states of coating activity for applied voltages less than about 0.5 volt. For these small applied voltages, activation and deactivation effects were absent and the voltage-current curves were symmetrical for either direction of current flow. From the slope of the voltage-current curves and the coating geometry the effective specific electrical conductivity could be computed.

The ac heater current for the nickel buttons was stabilized by a constant voltage transformer. For conductivity measurements the temperature was determined by measuring the thermal emf developed by one of the Mo−Ni thermocouples. It was found that the most sensitive method for setting the two nickel buttons to the same temperature was to adjust the current in the second heater until no thermal emf could be observed across the sample.

In measurements of the thermoelectric power the heater current was supplied by storage batteries. This was found necessary because the ac source was not sufficiently well regulated. The thermal emf across the sample was measured between the two nickel thermocouple leads using a Leeds and Northrup Type K-2 potentiometer with a galvanometer as a null indicator. The thermal emf of each Mo−Ni thermocouple was measured with the Type K-2 potentiometer.

The thermoelectric power, emf/ΔT, was determined from the slope of a plot of the thermal emf developed across the sample *versus* the difference in temperature, ΔT, of the two buttons. The average temperature of the sample was kept nearly constant in any given determination. A typical plot of the thermal emf across a sample *versus* ΔT is shown in Fig. 2.

Pulsed voltage-current characteristics were taken on many of the samples. The pulses were obtained from a square wave, hard tube, pulse generator with pulse length variable from 1 μsec to several hundred μsec.[6] A 3 μsec pulse at a repetition rate of 60 per second was used in most of these measurements. A Sylvania Model P-4 synchroscope was used for measuring the pulse voltage and a Tektronix Model 511-AD oscilloscope with a Type 121 preamplifier was used to measure the pulse current through the sample by means of suitable viewing resistors. Using the above equipment, pulse voltages up to 700 volts could be applied and pulse currents of from 1 ampere to 10^{-5} ampere could be measured.

[6] F. A. Horak, Technical Report No. 7. N7onr-202 T.O. 5 (University of Missouri, August, 1951), p. 4.

III. RESULTS AND CONCLUSIONS

The eight experimental tubes can be divided into three groups; a typical tube from each of these groups will be discussed in detail. Results obtained on the remaining tubes will be presented in the form of a table at the end of this section. The tubes were numbered in chronological order of their processing; however, tube No. 8 will be discussed first since the most complete set of observations were made on this sample.

A. Tube No. 8

This tube was constructed as indicated in the previous section, Fig. 1C. The 0.030-in. tantalum wire ring anode was introduced in order to make separate measurements of the external thermionic emission and to determine if part of the conduction current was being passed around the outside of the coating. The flat ends of the nickel buttons were completely covered with the carbonate; they were then pressed together to enable the electric field produced by the anode to reach the exposed coating surface between the nickel buttons. The coating was (BaSr)O on the standard nickel base metal.

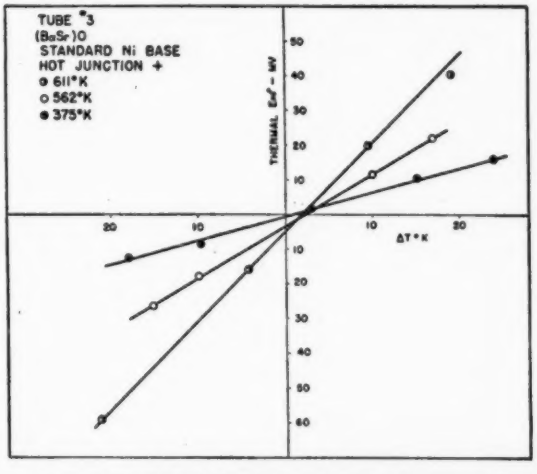
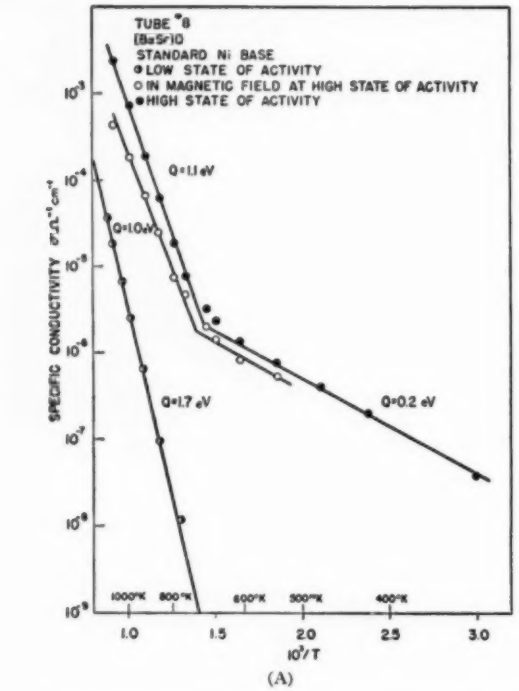
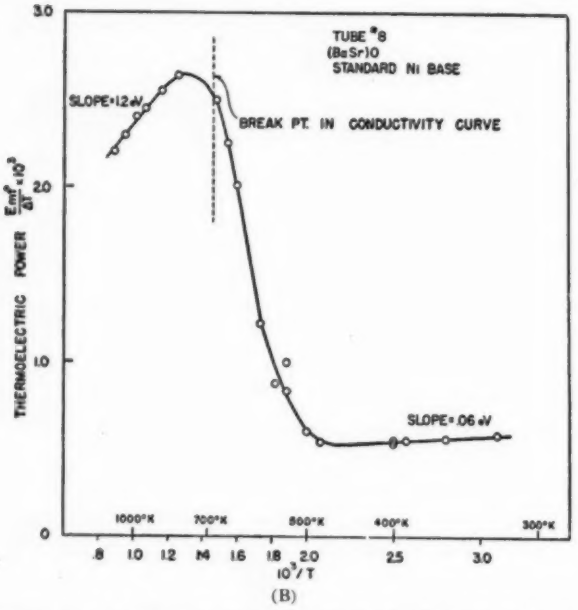


FIG. 2. A typical plot of thermal emf *vs* difference in temperature, (BaSr)O.

Conductivity measurements at two states of activity are shown in Fig. 3(A). In the low state of activity a straight line is observed over the entire temperature range in which measurements could be made; however, in the high state of activity two straight line sections are observed. The activation energies, Q, in the different states of activity are indicated on the figure. These results are typical of the conductivity observations made on all tubes and are in agreement with the results reported by Loosjes and Vink, except the region of the break in our conductivity curves appears at somewhat lower temperatures. In states of low activity the break was not always found; however, this is probably a consequence of our being unable to make measurements at temperatures sufficiently low to show its existence. The observed experimental curves are found to be quite similar to the theoretical conductivity curve of Hensley[3] based on the pore conduction hypothesis. These conductivity results may also be interpreted qualitatively using only semiconductor theory if two sets of energy levels are present to contribute electrons to the conduction process. However, quantitatively it does not seem likely that this model could account for the rather abrupt break in the curve which is observed. The one impurity level, simple semiconductor model can account for the behavior in one temperature region but not both.



(A)



(B)

FIG. 3. (A) Electrical conductivity of (BaSr)O *vs* reciprocal temperature. (B) Thermoelectric power of (BaSr)O *vs* reciprocal temperature.

If the conduction current is due to an electron gas in the pores, a lower conductivity would be expected if the sample were placed in a magnetic field. A much smaller change would be expected if the conduction is due to electronic conduction through the crystals. With a magnetic field of 5600 gauss applied perpendicular to the direction of current flow through the coating, the conductivity measurements were repeated. These values are represented by the open circles in Fig. 3(A). It was observed that the conductivity was reduced by a factor of up to six in the high temperature region and by a factor of two in the low temperature region. Thus the results are qualitatively in agreement with the pore conduction mechanism. If an appreciable fraction of the current is carried by electron emission around the sample a reduction of current in the magnetic field would also occur. An experiment to be described later seems to rule out this possibility.

Thermoelectric power measurements were taken with the sample in this high state of activity. A plot of the thermoelectric power *versus* reciprocal temperature is shown in Fig. 3(B). This plot is similar in shape and magnitude to Hensley's[3] theoretical curve for a porous semiconductor. The high temperature region had a slope of 1.2 ev which was in fair agreement with 1.1 ev obtained for the conductivity activation energy Q in the high temperature region. A rapid decrease of thermoelectric power occurred as the temperature was reduced from 700°K to 500°K. At still lower temperatures a slight rise in thermoelectric power was found; the latter region has a slope of 0.06 ev. This value does not agree favorably with the activation energy obtained from the conductivity curve in the low temperature region, $Q=0.24$ ev. The thermoelectric power measurements are in agreement with the predictions of the porous semiconductor theory on the following points: (1) The temperature dependence of the thermoelectric power in the high temperature region is essentially the same as the temperature dependence of the conductivity in this region. (2) The magnitude of the thermoelectric power, 2–2.5 mv/°, observed at high tempera-

tures is in fair agreement with the predictions of the porous semiconductor theory.[3] (3) The temperature dependence in the high temperature region is also the same as that of the thermionic emission from the oxide. (4) All samples, active and inactive, had a thermoelectric power in such a direction that the hot junction was positive over the entire temperature range. This is indicative that negative charge carriers are predominant in the electron gas at high temperatures and in the crystals at low temperatures.

The one impurity level, simple semiconductor theory predicts a linear dependence of the thermoelectric power and the reciprocal temperature.[3] Thus this model could account for either the high or the low temperature regions but not both. If a two impurity level model is used to account for the two slopes of the conductivity and the thermelectric power curves, both of these must be donor type levels to give an N type semiconductor at all temperatures. This model is not attractive as it also predicts a break in the thermionic emission curve which, as we will see later, is not observed.

In the low temperature region some disagreement is found between the slope of the thermoelectric power curve and the activation energy obtained from the conductivity curve. If the low temperature conductivity and thermoelectric power are due entirely to electronic conduction in the bulk oxide, then the two slopes should agree. It may be possible that measurements of the thermoelectric power at low temperatures are influenced by the mechanism which is predominant at high temperatures. Hensley[3] has shown that a leakage resistance in the tube greater by a factor of 10 than the sample resistance and in parallel with the sample, would short-circuit the thermoelectric power and cause the curve to have a considerably smaller slope in the low temperature region. It is possible that both of these effects could contribute to the disagreement between the activation energy obtained from the conductivity curve and the slope of the thermoelectric power curve in the low temperature region.

Dc voltage-current curves were taken at different temperatures, with applied voltages up to 6 volts. Figure 4(A) shows two such curves, one at 1027°K and the other at 533°K. The curve taken at 1027°K is typical of those above the break in the conductivity curve, and the one taken at 533°K is typical for those below the break. At temperatures above the break the curves appear concave toward the voltage axis, while at temperatures below the break the curves are convex toward the voltage axis.

Pulsed voltage-current measurements were also made at various temperatures. The pulse length used was 3μs, and up to 500 volts was applied, corresponding to fields up to 25,000 volts per cm. Figure 4(B) shows two pulsed voltage-current curves which were obtained, one at 982°K, the other at 652°K. The curve taken at 982°K is typical of those above the break in the conductivity curve, and the one at 652°K is typical for temperatures below the break. At 982°K the curve appears to be concave toward the voltage axis at low voltages and convex toward the voltage axis at higher voltages. The 652°K curve appears to be convex toward the voltage axis except in the low voltage region where it is nearly linear. No difference in the dc and pulsed values was found in the voltage range covered by both measurements.

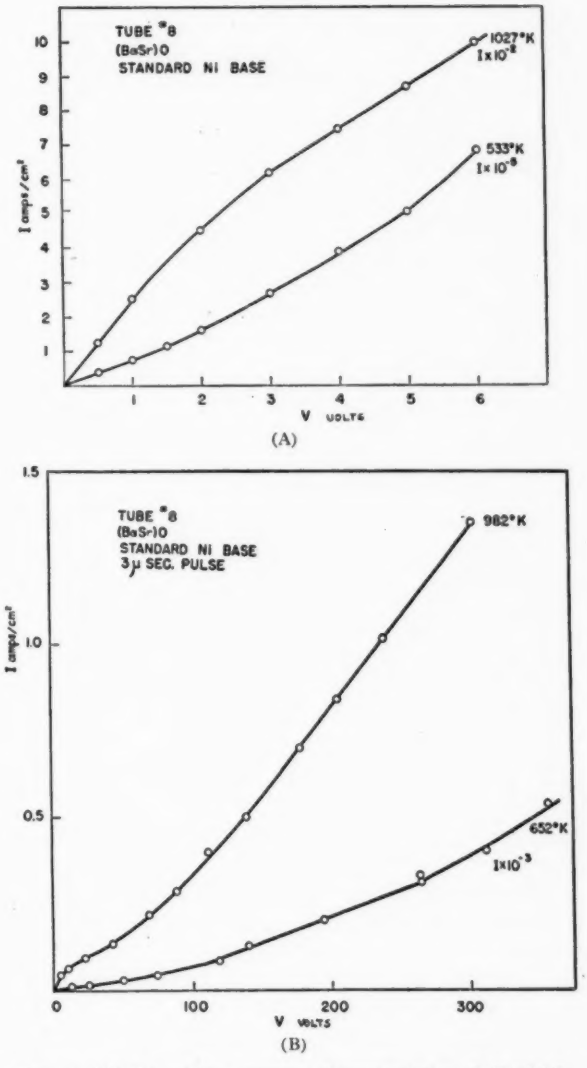


FIG. 4. (A) Dc voltage-current characteristics of (BaSr)O. (B) Pulsed voltage-current characteristics of (BaSr)O.

If, according to the pore conduction theory, the current is carried by an electron gas in the pores at high temperatures, the current-voltage dependence should resemble that of a diode. Accordingly a Schottky type plot ($\log I$ *vs* $V^{\frac{1}{2}}$) was made from the pulsed voltage-current data and is shown in Fig. 5. It is seen that straight lines result and the slopes of the lines vary in accord with the predictions of the Schottky equation. When these lines are extrapolated to zero voltage the zero field, internal, emission current is obtained. A

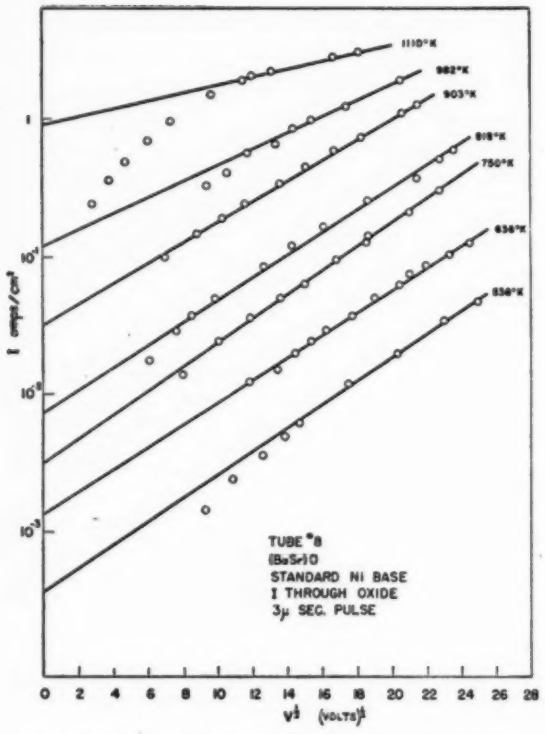


FIG. 5. A Schottky plot from pulsed voltage-current characteristics through (BaSr)O.

Richardson plot was constructed from these values of emission current and the temperature. Figure 6 shows this plot and indicates an internal work function of 1.1 ev. At temperatures below the break point in the conductivity curve, 800°K, the points on the Richardson plot are found to deviate from the straight line drawn through the high temprature points. This is to be expected since the current values below this temperature arise from a different mechanism.

The external thermionic work function was determined by drawing dc emission current from the exposed edge of the sample to the ring anode. Schottky plots made from this current and the applied anode voltage are shown in Fig. 7. From the extrapolated zero field emission currents a Richardson plot was constructed (Fig. 8). The external work function obtained from this plot was 1.1 ev, a value in agreement with the apparent work function obtained from the internal emission currents, with the activation energy obtained from the conductivity curve and with the slope of the thermoelectric power curve in the high temperature region. It is to be noted that the lowest point shown on the Richardson line occurs at a temperature below the break in the conductivity curve and below the point of deviation from the Richardson line in Fig. 6. These observations, substantiated by data on the other tubes of this group, cannot be explained by the two-energy level, semiconductor model, but are in agreement with the pore conduction hypothesis.

The shape of the voltage-current curves in the low voltage region have been interpreted by Loosjes and Vink.[2] The linear relationship at low temperatures was explained as being caused by the ohmic resistance of the coating, and the concave curvature toward the voltage axis at high temperatures was caused by a space charge effect in the pores. This behavior is seen in Fig. 4(B) for potentials below 50 volts. Above 50 volts the rapid increase of current with increasing applied voltage in the low temperature region may result from a Schottky effect or field emission of the pore electrons, giving rise to a current which masks the crystal conduction current which was predominant at low voltages. The voltage-current characteristics through the coating which appear as straight lines on the Schottky type plot (Fig. 5) have slopes very similar to the slopes of Schottky plots of the external thermionic emission from the coating to the ring anode (Fig. 7). These slopes are greater than predicted by theory for clean metal surfaces by a factor of about five. Other observers[7-9] have found the slopes of Schottky plots for oxide cathodes to be greater than theory predicts by a factor of between three and nine.

Since the possibility existed that part of the conduction current was carried around the coating by thermionic emission the following experiment was conducted. At the highest temperature at which measurements were made, 1100°K, 0.1 volt was applied across the coating producing an average internal field between the buttons of about 5 volts per cm. A positive potential was applied to the ring anode with respect to the midpoint of potential between the buttons and produced an external field at the oxide surface which could be varied from zero to 800 volts per cm. The current through the coating was observed as the anode voltage was increased. No change was detected in the current through the coating as the external field was varied. This indicates that a negligible amount of the conduction current passes around the edges of the coating.

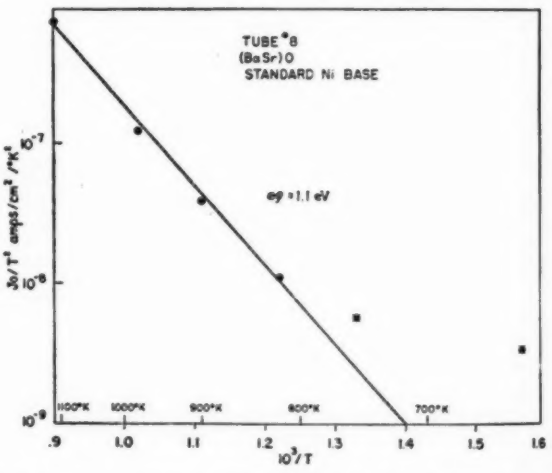


FIG. 6. A Richardson plot made from Schottky plot shown in Fig. 5.

[7] H. Y. Fan, J. Appl. Phys. **14**, 552 (1943).
[8] C. S. Hung, J. Appl. Phys. **21**, 37 (1950).
[9] I. L. Sparks, thesis, University of Missouri (1951).

B. Tube No. 4

This tube was constructed similar to Tube No. 8 except the ring anode was not present, but facilities were provided for introducing gases through glass break-off tips. The sample was (BaSr)O between base metals of standard nickel. After one hour of activation the tube was sealed off, the vacuum system and conductivity and thermoelectric power data were taken. Conductivity measurements in this state of activation are shown in Fig. 9(A), as open circles. Two straight line regions are apparent with the high temperature activation energy 1.6 ev and the low temperature activation energy 0.11 ev. Thermoelectric power measurements at this state of activation are shown in Fig. 9(B), as open circles. At higher temperatures the slope was about 1.3 ev, and as the temperature was reduced, the thermoelectric power dropped to a low value and remained nearly constant.

Activation was continued by passing direct current through the sample at 1150°K until no further change could be detected. Conductivity and thermoelectric power measurements were then repeated. The conductivity results are shown in Fig. 9(A) as solid dots. In the high temperature range the activation energy had decreased to 0.7 ev, while the low temperature activation energy remained at 0.11 ev. Thermoelectric power results are shown in Fig. 9(B) as solid dots. Only in the low temperature region could a slope be determined; this had a value of 0.04 ev.

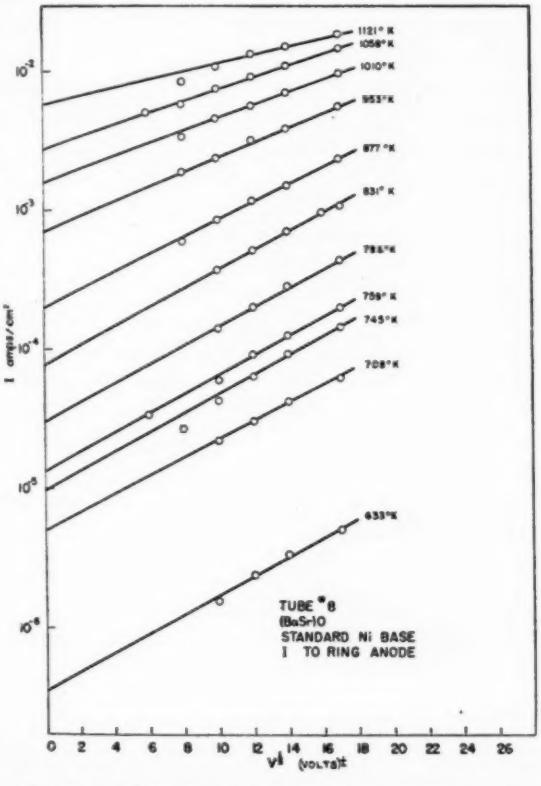


FIG. 7. A Schottky plot from emission current around the edges of the (BaSr)O sample.

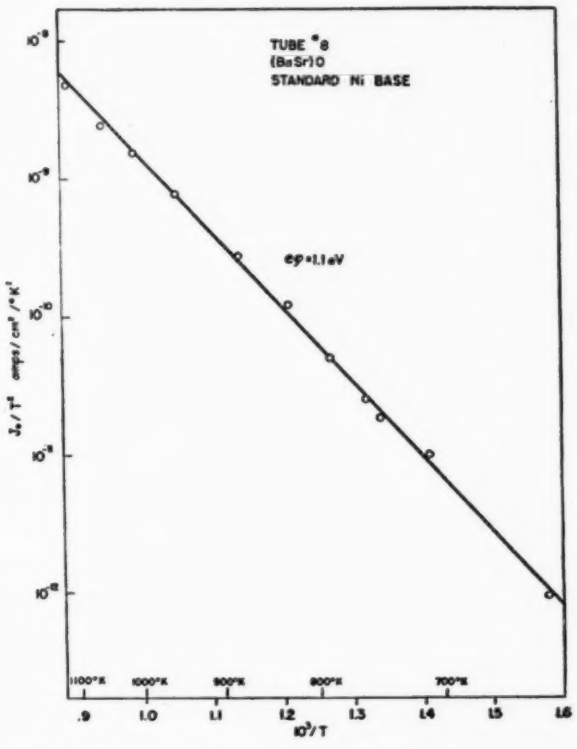


FIG. 8. A Richardson plot made from the Schottky plot in Fig. 7.

In order to obtain data on this sample in different states of activity and to determine, if possible, whether the low temperature conduction was due to surface conduction as suggested by Wright,[10] the sample was deactivated by introducing oxygen at room temperature. The tube was sealed at one of the crack-off tips to a vacuum system which was evacuated by a two-stage mercury diffusion pump. Two traps were located between the diffusion pump and the tube, one was cooled with a dry ice acetone mixture, while the other was cooled with liquid air. The vacuum system was pumped down, without baking, to a pressure of 10^{-7} mm of Hg. At this pressure the break-off tip was cracked by means of a Kovar rod placed inside the pumping lead, using an external magnet. Oxygen was obtained from a Linde Air Products Company cylinder and was filtered through a Koby activated charcoal filter before passing through the two cold traps.

A period of three months elapsed between the previous measurements on this tube and this oxygen experiment. After these data were recorded it was noted that the state of activity of the sample had changed in the three months that the tube was idle. Therefore the only reference to the activity before the introduction of oxygen, in the first deactivation experiment, was the conductivity at room temperature, not represented on Fig. 9(A).

[10] D. A. Wright, Phys. Rev. 82, 574 (1951).

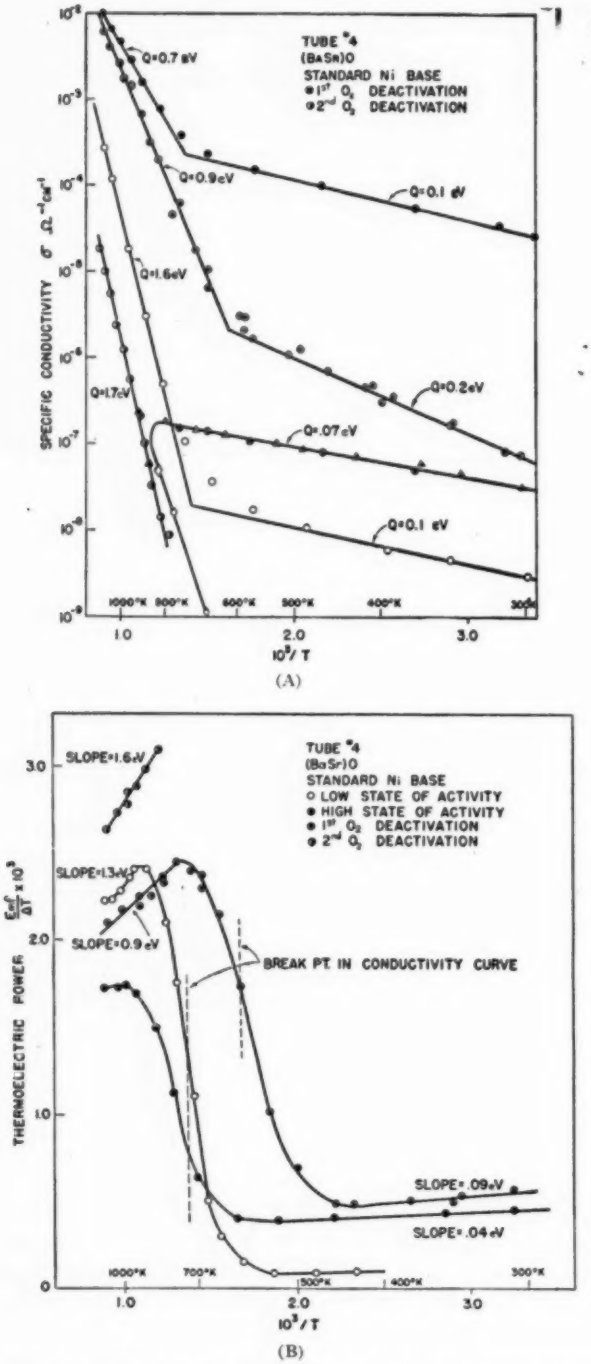


FIG. 9. (BaSr)O at various states of activation. Deactivation produced by oxygen treatment at room temperature. (A) Electrical conductivity. (B) Thermoelectric power.

Conductivity measurements were made at room temperature as the oxygen was introduced. Immediately after cracking the break-off tip and exposing the sample to the manifold vacuum of 10^{-7} of Hg, the conductivity reduced to two-thirds of its former value. When oxygen was introduced at a pressure of approximately 10^{-4} mm of Hg for a short time, only a slight further reduction of conductivity was observed. Oxygen was again introduced at a pressure of approximately 10^{-2} mm of mercury, and the conductivity fell to one-third of the value it had before cracking the break-off tip. Prolonged exposure to oxygen at pressures of 10^{-2} mm of mercury produced no further change in the conductivity at room temperature. The excess oxygen was then pumped out so that the pressure in the tube was less than 10^{-6} mm of mercury; no change in conductivity was observed during this re-evacuation. As the temperature of the sample was increased, the conductivity followed the curve shown by the △ points in Fig. 9(A). At approximately 850°K the conductivity fell rapidly, and at 900°K activation was observed. The temperature was then reduced and the new conductivity values were found to fall below the previous curve, giving the ◇ points on Fig. 9(A). Complete deactivation of the low temperature region is thus indicated. The temperature was then raised to 1100°K and current drawn until no further activation occurred. Conductivity and thermoelectric power measurements were made in this state of activity. The conductivity curve is represented by the ⊙ points in Fig. 9(A). The high temperature activation energy was 0.9 ev and the low temperature activation energy was 0.20 ev. Thermoelectric power measurements at this state of activity are represented by the ⊙ points in Fig. 9(B). The high temperature slope was found to be 0.9 ev and the low temperature region had a slope of 0.09 ev. Again close agreement is found between the value of the high temperature activation energy and the high temperature slope of the thermoelectric power curve, and the low temperature regions are not in good agreement.

Oxygen was again introduced at a pressure of 10^{-2} mm of mercury, with the sample at room temperature. The conductivity immediately fell to the same value observed when oxygen was introduced the first time. The tube was evacuated, and the temperature was increased in steps at intervals of 10 minutes. Conductivity results are indicated by the ⊗ points of Fig. 9(A); however, no change in conductivity with time during the 10-minute interval was noted until a temperature of 900°K was reached and at this point activation occurred. In order to produce a lower state of activity, oxygen was introduced again at a pressure of between 10^{-3} and 10^{-4} mm of mercury, and the temperature of the cathode was increased at 900°K for a few minutes. This resulted in a stable inactive state shown by the ◐ points in Fig. 9(A) obtained with the sample in vacuum. The activation energy for this state of activity was 1.7 ev. Thermoelectric power measurements are represented by the ◐ points in Fig. 9(B). The slope of the thermoelectric power plot is 1.6 ev in agreement with the activation energy obtained from the conductivity plot.

After these measurements were completed the sample was heated to 1150°K and reactivated by passing direct current in first one direction and then in the other direction. Polarization effects were noticed at first,

but as activation proceeded, the polarization effects disappeared. After activating for 15 hours, the sample was found to be in a state close to that obtained after the first oxygen deactivation, shown by the ⊙ curves in Figs. 9(A) and 9(B).

These activation and deactivation studies seem to indicate that the low temperature conductivity results from electronic conduction through the crystal grains. The initial deactivation that occurs when the sample is exposed to oxygen at room temperature probably results from a deactivation of the surface and the bulk a short distance inside the crystals. Since the diffusion rate of oxygen is undoubtedly low at room temperature, very little oxygen could diffuse into the bulk of the oxide crystals. This is borne out in Fig. 9(A) where the conductivity maintains a low slope region until a temperature of the order of 850°K is reached. Inward diffusion of the oxygen trapped at the surface becomes more rapid, and the low slope region is deactivated. This deactivation at 850°K could also result from the absorbed oxygen reacting with free barium on the surface of the crystals as suggested by Wright,[10] but this chemical reaction should take place at room temperature.

This experiment supports the pore conduction theory in that the low slope region extends to higher temperatures, about 850°K, after the oxygen treatment at room temperatures than before the treatment, about 650°K. Deactivation of the surface region would greatly reduce the thermionic emission at a given temperature, and if the high temperature conductivity is due to electron emission through the pores, the resulting apparent conductivity would be lowered. Thus one would be able to go to a higher temperature before pore conductivity becomes comparable to the bulk conductivity.

C. Tube No. 5

Since the electron gas in the pores is being considered as a possible conduction mechanism, it was thought desirable to construct a tube which could only pass current by the pore conduction mechanism. This tube was designed similar to the tubes mentioned previously except the coating was sprayed over the entire flat surface of each nickel button, and they were not pressed tightly together but the oxide surfaces were separated by about 0.3 mm. The only mechanism of passing current through the gap is by thermionic emission. Coatings were made of (BaSr)O and the base metal was standard nickel.

The sample was activated similar to the tubes previously described. Voltage-current curves were found to be symmetrical and nearly linear over the voltage range from 0 to 0.5 volt. Apparent conductivity measurements were made from the slope of the voltage-current curves through zero voltage and from the geometry. Figure 10(A) shows the conductivity plot that was obtained. Curvature is noticed at high temperatures with a straight line region appearing at the lower temperatures. A temperature dependence of this general shape was predicted by Loojes and Vink[2] for the conduction through the pores of alkaline earth oxides. The activation energy obtained from the slope of this curve was 1.8 ev, and from the pore conduction theory this value is approximately the work function.

Thermoelectric power measurements were made on this tube and the hot junction was again found to be positive. The experimental values scattered somewhat as indicated in Fig. 10(B), but the best straight line gave a slope of 1.9 ev. This slope is similar to the activation energy obtained from the conductivity curve in agreement with the pore conduction theory. Specific values of the thermoelectric power are seen to agree closely with values found in the high temperature region for the other samples which are pressed together.

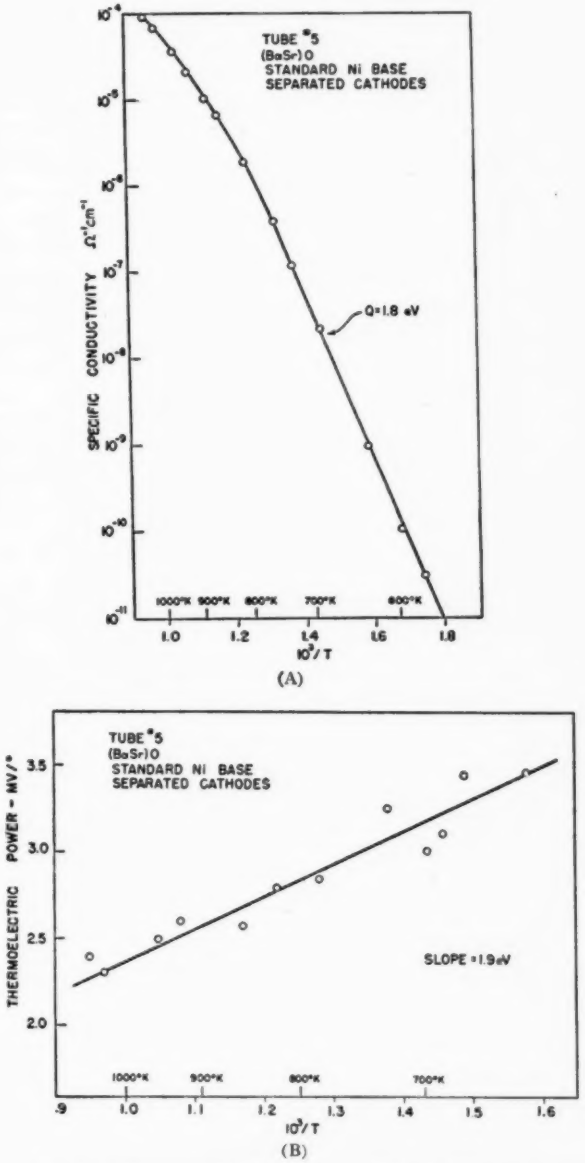


FIG. 10. Characteristics of separated (BaSr)O samples. (A) Apparent conductivity. (B) Thermoelectric power.

TABLE II. Summary of results on eight experimental tubes.

Tube No.	Coating	Base metal	Activation energy high temperature (ev)	Activation energy low temperature (ev)	Slope thermoelectric power high temperature (ev)	Slope thermoelectric power low temperature (ev)	Work function from $V-I$ through coating (ev)	Work function from $V-I$ to ring anode (ev)	Effective conductivity @ 1000°K (Ω^{-1} cm^{-1})	Thermoelectric power at 1000°K (mv/degree)
1	(BaSr)O	1001 Ni	1.7	...	1.6	...	...	...	1.2×10^{-5}	2.72
2	(BaSr)O	1001 Ni	1.6	0.7	1.4	...	...	...	1.1×10^{-5}	2.28
7	(BaSr)O	1001 Ni	1.7	1.0	1.8	...	1.6	1.8	7.0×10^{-6}	2.25
3	(BaSr)O	Std. Ni	0.9	0.22	0.7	0.09	1.0	...	2.0×10^{-3}	2.17
4	(BaSr)O	Std. Ni	1.6	0.11	1.3	...	...	...	4.1×10^{-5}	2.28
4	(BaSr)O	Std. Ni	0.7	0.13	...	0.04	...	...	5.0×10^{-3}	1.75
4	(BaSr)O	Std. Ni	1.7	...	1.6	...	...	...	1.9×10^{-6}	2.78
4	(BaSr)O	Std. Ni	0.9	0.20	0.9	0.09	...	...	2.5×10^{-3}	2.16
5	(BaSr)O	Std. Ni	1.8	...	1.9	...	...	...	7.0×10^{-5}	2.36
8	(BaSr)O	Std. Ni	1.1	0.24	1.2	0.06	1.1	1.1	8.0×10^{-4}	2.37
6	BaO	Std. Ni	1.0	0.44	0.9	...	...	...	1.1×10^{-3}	2.12
6	BaO	Std. Ni	0.9	0.42	0.9	...	1.0	...	2.0×10^{-3}	2.06

A summary of the results obtained for the different tubes may be found in Table II. According to the pore conduction theory, the ratio of the electrical conductivity to emission current density σ/J should be about 3×10^{-3} cm/Ω amp at 1000°K. Table III lists the σ/J results obtained in this investigation. For the internal thermionic emission current the theoretical value seems to be reasonable, but the external emission current values appear to be large. Since the geometry was not ideal and the external emitting area uncertain, the external σ/J values could be expected to be in error. However, it is difficult to explain why these ratios differ by such a large factor.

IV. SUMMARY

No interface effects were found in comparing the results obtained on samples made on the standard nickel base metal and on the 1001 pure nickel base. The only difference observed on samples made on the pure nickel base was their relative low activity and the difficulty encountered in activating. This is in agreement with the results of other observers.

No significant difference was found in the properties of BaO as compared to (BaSr)O. Activation was slightly more difficult and the low temperature activation energy, 0.4 ev, was larger than that observed for (BaSr)O. This disagrees with Loosjes[11] who measured the energy distributions of electrons emitted from BaO and (BaSr)O and found them to be different. He concluded that the conduction mechanisms were different and that BaO was not a pore conductor.

TABLE III.

Tube no.	σ/J cm/Ω amp internal	σ/J cm/Ω amp external
3	8×10^{-3}	...
6	1×10^{-2}	...
7	3×10^{-2}	7×10^{-1}
8	4×10^{-3}	4×10^{-1}

Specific values of the thermoelectric power in the high temperature region appear to be related to the state of activity of the sample. As the activation energy and slope of the thermoelectric power curves reduce with increasing activity of the sample, the specific value of the thermoelectric power at high temperatures likewise decreases. This is in accord with the predictions of the porous semiconductor theory.[3]

The simple one- or two-level, semiconductor model does not seem to adequately explain these various experimental results which are in agreement with the pore conduction theory. The major disagreement with this latter theory is the slope of the low temperature region of the thermoelectric power curve which does not agree well with the activation energy in the same temperature range. A possible explanation for this discrepancy was discussed.

The author wishes to express his appreciation to Professor A. S. Eisenstein, whose suggestions led to the undertaking of the problem and under whose guidance the investigation was completed.

To the Office of Naval Research and the University of Missouri Research Council the author is indebted for the support which made this work possible.

[11] R. Loosjes, private communication.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 10 OCTOBER, 1952

A Method of Solving the Wave Equation in a Region of Rapidly Varying Complex Refractive Index*

J. J. GIBBONS AND R. L. SCHRAG
Ionosphere Research Laboratory, The Pennsylvania State College, State College, Pennsylvania
(Received April 21, 1952)

A process is described whereby the complex wave equation is transformed into an integro-differential equation in a single real dependent variable. This makes possible a technique for obtaining wave functions, which is believed to offer definite advantages in some instances where the more standard methods are difficult to handle, e.g., in regions where the complex index is varying rapidly. The index may be entered in a tabulated form, and its exact analytical expression is not required.

The technique is then demonstrated by computing the reflection coefficient associated with an extremely critical region, that which one encounters at the "fourth reflection condition," where the refractive index goes through a sharp peak (see reference 3).

INTRODUCTION

INVESTIGATIONS into wave propagation theory for ionized media frequently require solutions to the wave equation[1]

$$\pi''+K_0^2n^2\pi=0, \tag{1}$$

where the variable refractive index n is a complex quantity. Perhaps the most direct method of attack is to separate the wave equation into its real and imaginary parts, thus obtaining two equations that are coupled in the sense that each involves the real and imaginary components of π. Their simultaneous solution is then required. Well-behaved solutions by this procedure are sometimes difficult to obtain, particularly when the integration is attempted through regions of rapidly varying index.

The process that we are about to consider transforms the complex wave equation into an integro-differential equation in one real dependent variable, thus avoiding the necessity of solving two simultaneous differential equations. The solution of this integro-differential equation yields one of the two wave functions directly. The second solution is obtained from the first by direct integration. The technique appears to yield workable wave functions in critical regions where other known procedures fail.

THE METHOD

Let us consider a wave function expressed in the form

$$\pi=ze^{jQ}=\exp\left(\int e[(z'/z)+jQ']dx\right). \tag{2}$$

Substitute this expression into (1) and separate into real and imaginary parts to obtain

$$(z''/z)-Q'^2=R \tag{3}$$

and

$$Q''+2(z'Q'/z)=I, \tag{4}$$

where R and I are defined by

$$R+jI=-K_0^2n^2. \tag{5}$$

After multiplying Eq. (4) by z^2, it may be integrated directly to obtain

$$Q'=(1/z^2)\int Iz^2dx, \tag{6}$$

which, substituted into (3), yields

$$(z''/z)-\left[(1/z^2)\int Iz^2dx\right]^2=R. \tag{7}$$

A direct integration of Eq. (7) produces the amplitude function and the first derivative of the phase function of the first wave solution. An additional integration on Q' then completely determines the wave function π. We observe that for the case where the refractive index is real, $I=0$, and Eq. (7) reduces to the original wave equation.

The second of the two required wave functions in any region may be found by repeating the integration a second time, using different initial conditions, or by employing Abel's integral formula.[2] For our purpose, the latter may be written

$$\pi_2=\pi_1\int dx/\pi_1^2, \tag{8}$$

or, in terms of our present notation,

$$\pi_2=\pi_1\left[\int(\cos 2Q/z^2)dx-j\int(\sin 2Q/z^2)dx\right]. \tag{9}$$

In general, the π-functions will be involved in a boundary value problem which requires that they be matched smoothly onto other wave functions. For this purpose,

* The research reported in this paper has been supported by the Geophysics Research Division of the Air Force Cambridge Research Center under Contract AF19(122)-44 and by the National Science Foundation under Grant NSF-066.

[1] O. E. H. Rydbeck, Trans. Chalmers University Gothenburg Nr. 34, 1944.

[2] R. P. Agnew, *Differential Equations* (McGraw-Hill Book Company, Inc., New York, 1942).

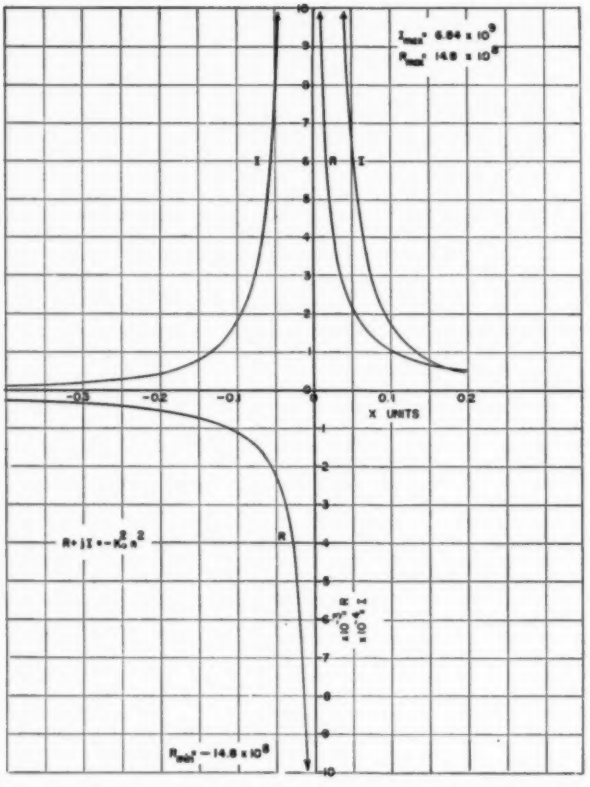


FIG. 1. Plots of R and I defined by formulas (12) and (13) for the non-W.K.B. interval. (This interval contains the "fourth reflection condition.")

it is helpful to realize that

$$\pi_2' = (1/\pi_1)[1+\pi_1'\pi_2], \tag{10}$$

where

$$\pi_1' = \pi_1[(z'/z)+jQ']. \tag{11}$$

The use of these formulas provides an easy and accurate method for obtaining the first derivatives.

The technique by which the above method is employed will now be illustrated by an example.

EXAMPLE

In order to illustrate the applicability of the above method of obtaining wave solutions under extreme conditions, a solution has been carried through an ionospheric region containing a sharp peak in its

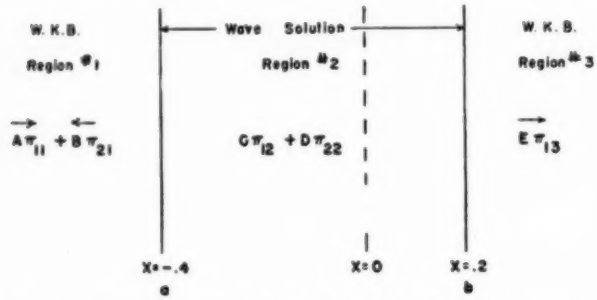


FIG. 2. The configuration for the boundary value problem. ("Region 2" is the non-W.K.B. interval corresponding to Fig. 1.)

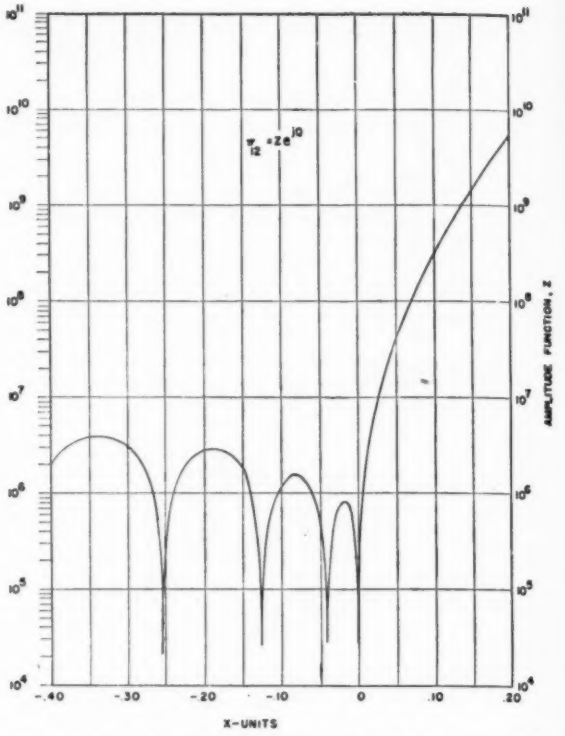


FIG. 3. The amplitude functions z of the first wave solution in "Region 2."

refractive index. The condition giving rise to the corresponding zero-collision pole is discussed in detail by S. K. Mitra,[3] and is commonly called the "fourth reflection condition." Since it is, in general, impossible to construct the required wave solution from elementary analytic functions,[4] the process necessarily consists of numerical integrations. To make such computations possible, the effect of a small collisional frequency has been included, thus making the index components finite at the peak and also eliminating the triviality of having I vanish everywhere.

Throughout the range of interest, i.e., where wave solutions are required, the R and I functions can be assumed to be given by

$$R = -K_0^2[1+K_1x/(\eta+x^2)], \tag{12}$$

and

$$I = -K_0^2K_2/(\delta+x^2), \tag{13}$$

where x is the height variable measured from the peak in the index, being positive upwards and negative downwards. For our purposes, the wave frequency as well as the characteristics of the medium will be defined by the following parameters:

$$K_0=\pi, \quad K_1=-11.0, \quad K_2=-1.8\times10^{-7},$$
$$\delta=2.6\times10^{-16}, \quad \eta=13.5\times10^{-16}.$$

The R and I functions for these parameters are shown plotted in Fig. 1. It is impossible to display their

[3] S. K. Mitra, *The Upper Atmosphere* (The Royal Asiatic Society of Bengal, 1947).

[4] J. J. Gibbons and R. J. Nertney, J. Geophys. Research 56, 355 (1951).

behavior in the immediate vicinity of $x=0$ without greatly expanding the scales. The peak values are, however, indicated in the figure.

Standard considerations show that for a region defined by (12) and (13), it is permissible to employ W.K.B. solutions provided $|x|>0.2$. The problem of determining the reflection coefficient associated with the critical region (see Fig. 2) consists of computing the wave functions, π_{12} and π_{22}, in the non-W.K.B. region, and matching these functions smoothly onto the W.K.B. solutions at the respective boundaries a and b. The W.K.B. solutions are given by

$$\left.\begin{aligned}\pi_{11}&=[1/(K_0n)^{\frac{1}{4}}]\exp\left(-jK_0\int_a^x ndx\right),\\ \pi_{21}&=[1/(K_0n)^{\frac{1}{4}}]\exp\left(jK_0\int_a^x ndx\right),\\ \pi_{13}&=[1/(K_0n)^{\frac{1}{4}}]\exp\left(-jK_0\int_b^x ndx\right),\end{aligned}\right\} \quad (14)$$

where π_{11} and π_{13} represent up-going waves and π_{21} is a down-coming wave. This treatment, therefore, assumes that the wave transmitted into Region No. 3 will not be reflected from a higher level.

In carrying out the matching problem at the boundaries a and b, we simply require that

$$\left.\begin{aligned}A\pi_{11a}+B\pi_{21a}&=C\pi_{12a}+D\pi_{22a}\\ A\pi'_{11a}+B\pi'_{21a}&=C\pi'_{12a}+D\pi'_{22a}\\ C\pi_{12b}+D\pi_{22b}&=E\pi_{13b}\\ C\pi'_{12b}+D\pi'_{22b}&=E\pi'_{13b}\end{aligned}\right\}, \quad (15)$$

where the primes represent derivatives with respect to the variable x. After all the π-functions as well as their first derivatives have been determined at their respective boundaries, the system of Eqs. (15) may be solved to obtain the reflection coefficient,

$$\text{Refl. coeff.}=|B/A|.$$

There remains the determination of the wave functions in Region No. 2. As outlined previously, the first solution is found by solving Eq. (7). This was accomplished on an I.B.M.-C.P.C. computor by a numerical integration process started at $x=0$ and carried in both directions. The assumed initial conditions were $z_0=10^5$ and $z_0'=10$. Figure 3 shows a plot of the resulting amplitude function z. It will be noted that this function increases monotonically for positive values of x. This is characteristic of Eq. (7) in a region where R is always positive and when the initial value of z, as well as its first derivative, is taken as positive. The peculiar cycloidal form of the amplitude function for negative values of x is likewise believed to be characteristic of Eq. (7) when R is negative.

As a by-product of the integration on Eq. (7), one has Q', the derivative of the phase factor. This was integrated to obtain the curve in Fig. 4. The phase factor is seen to change rapidly by approximately 3 radians for each "well" in the z function. We thus have completely determined the first wave function in Region No. 2,

$$\pi_{12}=ze^{jQ}.$$

The second function π_{22} was computed from Eq. (9) only at the boundaries a and b, and thus cannot be shown plotted as a function of x. In determining the

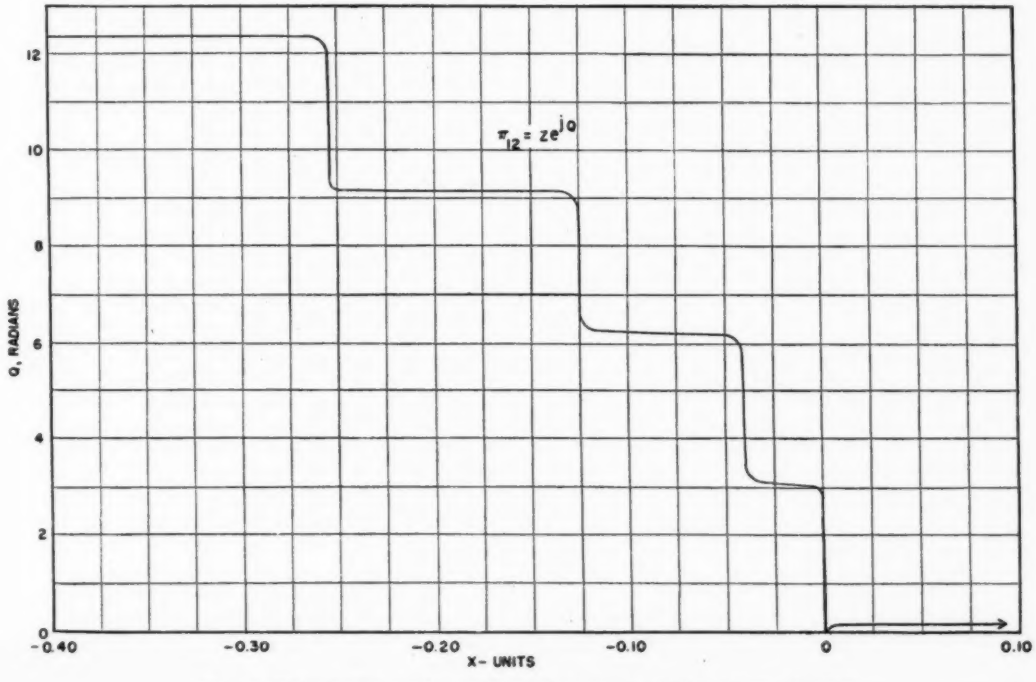


FIG. 4. The phase function Q of the first wave solution in "Region 2."

derivatives of all the wave functions, it was necessary to use Eqs. (14), (10), and (11). The numerical details are not included here.

The final step in this example consists of solving the system of Eqs. (15) for the ratio B/A. The result

$$|B/A| = 0.983$$

was obtained from the boundary configuration of Fig. 2, and checked by moving the upper boundary b from 0.2 to 0.1. These results show that a wave, with the assumed free space propagation constant, is virtually totally reflected when propagated into this critical region from below.

CONCLUSIONS

The method of wave solution described in this paper applies to the uncoupled wave equation with complex coefficients. It appears to have special advantages when these coefficients vary in a manner that causes the intermediate functions[4] of some methods to oscillate wildly. As illustrated in the example, the present method will, in general, also produce oscillatory amplitude functions with rather sharp "wells," which require special care. However, the oscillations proved to be few in number for even the extreme example considered.

Meeting of the Operations Research Society

The first regular meeting of the Operations Research Society, which was formed by a group of leading scientists in May, 1952, will be held at the National Bureau of Standards in Washington, D. C., on November 17th and 18th.

The Society, whose purpose is to promote the application of scientific methods to business and military problems, will hear and discuss papers involving industrial applications of Operations Research; waiting lines; inventories and the use of probability techniques in industry; and the newly developed Monte Carlo method for handling problems which cannot be solved with ordinary equations. There will also be a research panel on the solution of new problems and the development of new methods which have arisen in industry in the last year.

Professor Philip Morse, of the Massachusetts Institute of Technology, is President of the Society; the membership includes experts from such diverse disciplines as physics, chemistry, biology, mathematics, economics, statistics, etc.

Non-members are invited to attend the meeting and to register at the official headquarters at the Wardman Park Hotel, Washington, D. C.

Further information about the meeting may be obtained from the Secretary of the Society, Mr. John B. Lathrop, 30 Memorial Drive, Cambridge, Massachusetts.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 10 OCTOBER, 1952

A Note on the Approach of Narrow Band Noise after a Nonlinear Device to a Normal Probability Density

GEORGE R. ARTHUR*
Dunham Laboratory, Yale University, New Haven, Connecticut
(Received April 30, 1952)

The output of a low pass filter preceded by a nonlinear device excited by a noise source is given in terms of an integral equation. The first three central moments of the probability density of the output signal are obtained from the integral equation and the approach of these moments to those of a Gaussian or normal density is demonstrated.

CONSIDERABLE material has been written on the statistics of noise which has been passed through a nonlinear device. However, if the output of the nonlinear device is passed through a low pass filter, little is known of the exact statistical nature of that filter output, unless the output band is very narrow. In this case, the probability density of the output approaches a normal law although no good general proof of this is yet available. For special cases, the approach to a normal law may be demonstrated. Often for particular applications, it becomes necessary to know quantitatively just how narrow the output band must be for the approximation by the normal law to be a good one.

In an excellent article several years ago, Kac and Siegert[1] have shown that the characteristic function of the output of a low pass filter which is preceded by a square law device, a band limiting filter and a white noise source is

$$\Phi(u)=\frac{1}{\prod_{j=1}^{N}(1-iu\lambda_j)} \tag{1}$$

with the λ_j determined from the integral equation

$$\int_{-\infty}^{\infty} dt f(t)[K(s)]^{\frac{1}{2}}p(s-t)[K(t)]^{\frac{1}{2}}=\lambda f(s) \tag{2}$$

where the $K(t)$ is the weighting function of the low pass filter and the $p(t)$ the autocorrelation of the envelope portion output of the band limiting filter. The number of eigenvalues, N, may or may now be infinite, depending on the nature of the kernel.

If now the log of the characteristic function is considered, one has:

$$\log\Phi(u)=-[\log(1-iu\lambda_1)+\log(1-iu\lambda_2)+\cdots+\log(1-iu\lambda_N)]. \tag{3}$$

Using the expansion for $\log(1+z)$ Eq. (3) becomes

$$\log\Phi(u)=\sum_1^N \lambda_j(iu)+\frac{1}{2}\sum_1 \lambda_j^2(iu)^2+\frac{1}{3}\sum_1^N \lambda_j^3(iu)^3+\cdots+\frac{1}{N}\sum_1^N \lambda_j^N(iu)^N. \tag{4}$$

If the expansion of the log of the characteristic function is compared to the usual expansion of this function in terms of the semi-invariants of the probability density[2] corresponding to $\Phi(u)$, one has:

$$\sum_1^N \lambda_j=\chi_1=m$$

$$\sum_1^N \lambda_j^2=\chi_2=\sigma^2 \tag{5}$$

$$2\sum_1^N \lambda_j^3=\chi_3=\mu_3$$

and so on.

It is now possible to evaluate these coefficients in terms of the iterated kernels of the integral equation (2) and from these obtain a quantitative result for the approximation to the normal law. In other words, calling the kernel of Eq. (2), $R(x, z)$

$$\int_{-\infty}^{\infty} dx R(x, x)=\sum\lambda_j$$

$$\int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy R(x, y)R(y, x)=\sum\lambda_j^2 \tag{6}$$

$$\int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dz R(x, y)R(y, z)R(z,x)=\sum\lambda_j^3$$

and so on, with the only restriction being that the kernel be symmetric.[3]

It is possible in several cases to evaluate the first three semi-invariants in terms of the various weighting functions of the system in question with the results dependent on the ratio of low pass band width to noise bandwidth and thus clearly illustrate the approach to a normal density.

Since for a normal law only the first two semi-invariants exist and these are also the mean and

* Now Project Engineer, Sperry Gyroscope Company, Great Neck, New York.

[1] M. Kac and A. Siegert, J. Appl. Phys. 18, 383 (1947).

[2] Cramer, H., *Mathematical Methods of Statistics* (Princeton University Press, Princeton, 1946).

[3] H. Margenau and G. M. Murphy, *The Mathematics of Physics and Chemistry* (D. Van Nostrand Company, Inc., New York, 1943).

variance, respectively, the effect of the third semi-invariant or third central moment which introduces a skewness to the density function must become negligible to obtain a good approximation.

If the first three semi-invariants are considered as the dominant ones, the characteristic function is

$$\Phi(u)=\exp\left[-\frac{1}{2}\sigma^2u^2-i\left(um+\frac{\mu_3}{6}u^3\right)\right]. \tag{7}$$

The magnitude term here is normal and if the phase term remains almost linear over the range of u where the $|\Phi(u)|\gtrsim e^{-2}$, the Fourier transform or probability density is nearly normal with mean m and standard deviation, σ. Thus the value of interest is $u^2=4/\sigma^2$ with m normalized to one.

For this, the phase term is

$$\text{Arg }\Phi(u)=\exp\left[-i\frac{2}{\sigma}\left(1+\frac{2}{3}\frac{\mu_3}{\sigma^2}\right)\right]. \tag{8}$$

Thus if the ratio $\frac{2}{3}(\mu_3/\sigma^2)$ is less than say 5 percent, a good approximation to the normal law is obtained.

For the case considered by Kac and Siegert

$$\begin{aligned} K(t)&=\beta e^{-\beta t} \quad t>0\\ &=0 \qquad\quad t<0 \end{aligned} \tag{9}$$

and

$$p(l)=e^{-\alpha|l|}.$$

This means the low pass device is an integrator with time constant $1/\beta$ and the effect of the band limiting device is characterized by a frequency function

$$B(f)=\frac{\text{constant}}{1+\dfrac{2\pi if}{a}}. \tag{10}$$

The noise band width is taken as 2α and the output band as β. For this case:

$$\begin{aligned} \sum\lambda_j&=m=1\\ \sum\lambda_j^2&=\sigma^2=\frac{1}{\dfrac{2\alpha}{\beta}+1}\rightarrow\frac{\beta}{2\alpha}\\ \sum\lambda_j^3&=\frac{\mu_3}{2}\rightarrow 2\left(\frac{\beta}{2\alpha}\right)^2. \end{aligned} \tag{11}$$

A plot of $\frac{2}{3}(\mu_3/\sigma^2)$ *vs* $2\alpha/\beta$ falls below 5 percent at $2\alpha/\beta$ equal to 50. Thus for noise to low pass band ratios beyond this the approach to a normal law is quite close.

If Gaussian filters are assumed throughout with low pass band taken as σ_l and noise band as $2\sigma_N$ and:

$$\begin{aligned} K(l)&=\pi^{\frac{1}{2}}\sigma_l\exp(-\pi^2\sigma_l^2l^2)\\ p(l)&=\exp(-\pi^2\sigma_N^2l^2). \end{aligned} \tag{12}$$

This case yields:

$$\begin{aligned} \sigma^2&\rightarrow\frac{\sigma_l}{\sqrt{2}\sigma_N}\\ \frac{\mu_3}{2}&\rightarrow\frac{\sigma_l^2}{3\sqrt{2}\sigma_N^2} \end{aligned} \tag{13}$$

with $\frac{2}{3}(\mu_3/\sigma^2)$ below 5 percent for $2\sigma_N/\sigma_l$ about twenty. The order of magnitude is about the same as the previous case, the difference probably due to slightly different ways of defining band width. This method of moments seems especially useful and has been also used with slight modification to obtain quantitative results for a frequency discriminator with random signal impressed.

The Electronic Properties of Barium Sulfide*

WALTER GRATTIDGE† AND HAROLD JOHN
Department of Physics, University of Missouri, Columbia, Missouri
(Received April 24, 1952)

Barium sulfide has been prepared in a form suitable as a cathode emitter as well as in the form of compressed pills. It has a work function of 2.1 ev with a wide variation possible in activation. The most active materials had an emission at temperatures of 900°K and greater comparable to that from pure BaO. Pure nickel and platinum only were used as base metals. An equimolar mixture of BaS–SrS had a work function of 2.6 ev, and x-ray analysis indicated solid solution formation. The electrical conductivity of BaS was found to vary between wide limits depending on the degree of activation and the denseness of the specimen but was much less than for BaO. At 1000°K the specific conductivity of a compressed pill was 10^{-5} ohm^{-1} cm^{-1}. The thermoelectric power was found to be constant with temperature although it decreased with increase in the activity of the specimen. The warmer electrode was always positive. Pure BaS had a thermoelectric power of 2.5 mv per degree between 970° and 1270°K. The effect of added impurities of iron on the electrical characteristics was studied. By mass-spectrometer studies, the evaporation rate of BaS was found to be much less than that of BaO, indicating a vapor pressure of the sulfide about 100 times less than that of the oxide. The heat of sublimation was found to be 104 kcal/mole.

INTRODUCTION

FOR many years there has been considerable interest in the physical properties of the alkaline earth oxides and in their ability to emit electrons by heating, illumination, or by electron bombardment. Apart from luminescence studies there has been little systematic investigation of the electronic properties of the alkaline earth sulfides though the deleterious effects of sulfides in oxides have been studied.[1] The present investigation was initiated to produce barium sulfide, which has a molecular structure and chemical properties similar to barium oxide, and to investigate certain of its electrical characteristics.

PROPERTIES AND CHEMICAL PREPARATION

Barium sulfide is a hard white substance which is unstable in air gradually oxidizing to a yellowish color with liberation of hydrogen sulfide. It crystallizes as face-centered cubic with the edge-length of the fundamental cube 6.37A, about 14 percent larger than that of BaO whose lattice constant is 5.53A. According to Muller,[2] BaS forms mixed crystals with CaS but with SrS it is isomorphous. We have prepared a 50–50 (molar percent) mixed sulfide of barium and strontium from an equimolar mixture of the nitrates and subsequent reduction of the sulfate. X-ray powder pattern measurements indicated solid solution formation, the unit cell lengths being 6.368A for BaS, 6.20A for (BaSr)S, and 6.010A for SrS.

Several methods of preparation of BaS are indicated in the literature.[3,4] Two of these methods have been used, and their relative merits will be indicated.

(*a*) *From Barium Hydrosulfide.*—Hydrogen sulfide‡ is passed for 2 to 3 hours through a hot saturated solution of barium hydroxide.§ The solution is filtered and cooled. The addition of pure ethyl alcohol precipitates barium hydrosulfide as a hydrated white crystalline material which on long standing in air gradually turns yellow.

Barium sulfide can then be prepared from the hydrosulfide by heating in a stream of "forming" gas (92 percent nitrogen, 8 percent hydrogen) to a temperature of 800°C for several minutes.

(*b*) *From Barium Sulfate.* Barium sulfate is precipitated from a hot saturated solution of barium nitrate|| using sulfuric acid. This is a white precipitate which is stable in air. Barium sulfide is prepared by reducing the sulfate in platinum boats in a hydrogen atmosphere at 1000°C, care being taken to cool the specimen in the hydrogen stream.

The sulfide was always made into final usable form (e.g., sprayed on cathodes or pressed into pill form) immediately after preparation and stored under vacuum until required.

Spectrographic analyses were made of the various samples of BaS prepared. Apart from the initial preparation of BaS, the analyses indicate that both methods of preparation yield material of purity comparable with the barium carbonate used in the formation of barium

* Supported in part by the ONR.
† Now at the Research Laboratory for Electronics, Massachusetts Institute of Technology.
[1] H. A. Stahl, Appl. Sci. Res. **B1**, 397 (1950).
[2] W. Muller, Zentr. Mineral., Geol. **1**, 179 (1900).
[3] Roland Ward, *Preparation and Characteristics of Solid Luminescent Materials* (John Wiley and Sons, Inc., New York, 1948), Chapter II.
[4] L. Gmelin, *Handbuch der Anorganischen Chemie* (Verlag Chemie, Berlin, 1931), eighth edition, Vol. 30, p. 251.
‡ For the initial preparations, H_2S was generated by adding hydrochloric acid to ferrous sulfide. For later preparations, a cylinder of commercial H_2S (99.9 percent purity) obtained from the Matheson Company was used. Purification of H_2S was necessary in the initial preparations and was accomplished by bubbling the gas through two water columns.
§ The Baker and Adamson $Ba(OH)_2$ used indicated the following metallic impurities—Ca and alkali salts 0.2 percent, Fe 0.001 percent, and heavy metals 0.0005 percent.
|| Mallinckrodt, C. P. $Ba(NO_3)_2$ was used with indicated maxima for metallic impurities as follows: Ca and alkali salts 0.04 percent, Ca and Sr 0.02 percent, Fe 0.0002 percent and heavy metals (as Pb) 0.0005 percent.

TABLE I. Thermionic emission results on different cathode materials.

I	II	III	IV	V	VI	VII	VIII	IX
Material	Method of pareparation	Cathode No.	Anode	Cathode	i_0 at 1000°K amp/sq cm	$e\Phi$ (dc) ev	A (dc) amps/cm²/°K²	$e\Phi$ (pulse) ev
BaS	Method (*a*) from $Ba(SH_2)$ external conversion	I	Ta	1001 Ni	3.7×10^{-2}	2.2	4.4×10^{3}	
		II	Ta	1001 Ni	3.4×10^{-2}	2.0	29	
		III	Ta	1001 Ni	5×10^{-6}	2.0	0.06	
BaS	Method (*a*) $Ba(SH)_2$ applied	IV	Ta	1001 Ni	3.3×10^{-5}	2.15	2.4	
BaS	Method (*b*) from $BaSO_4$ external conversion	V	Ta	1001 Ni	1.6×10^{-5}	2.2	3.0	
		VI(*a*)	Pt	Pt	4.9×10^{-5}	2.1	2.4	3.0
		VI(*b*)	Pt	Pt	1.9×10^{-5}	2.0	0.34	2.1
		VII(*a*)	Pt	1001 Ni	5.0×10^{-4}	2.0	6.0	2.0
		VII(*b*)	Pt	1001 Ni	7.0×10^{-3}	1.8	2.2	2.4
(BaSr)S 50–50	Method (*b*)		Ta	1001 Ni	1.9×10^{-7}	2.6	2.0	2.68
BaO		A	Ta	1001 Ni	2×10^{-2}	1.37	0.122	
		B	Ta	1001 Ni	2×10^{-4}	1.79	0.22	
		C	Ta	1001 Ni	1×10^{-3}	1.5	0.05	

oxide cathodes for thermionic emission. These analyses are given in Appendix A.

(1) Thermionic Emission

Thermionic emission from barium sulfide was first investigated by O. W. Richardson.[5] In addition to electrons he also found evidence of positive ion emission. More recently Fox and Bailey[6] (using an electron microscope) made a comparative study of various materials as thermionic emitters. They found BaS to be a better emitter than SrO but not as good as BaO at a temperature of 1100°K.

The present emission studies have been made on barium sulfide as a cathode emitter in a diode with plane parallel geometry. Since the temperatures for emission from the sulfide are higher than from the oxide, a special tantalum cathode button was developed which could be faced with a disk of pure nickel¶ or platinum¶ as base for the sulfide. A Pt–Pt 10 percent Rh thermocouple was attached to the cathode for the purpose of temperature measurements. The cathode was prepared as a suspension of the sulfide or hydrosulfide in absolute ethyl alcohol and was applied to the flat cathode base by spraying or painting. Coating weights were in the range 12–15 mg per sq cm. The flat anode was made of 0.003-in. tantalum, and in tubes used in the latter part of the study this was faced with a 0.003 in. thick disk of platinum. Radiating fins of tantalum enabled increased anode heat dissipation. The form of the diodes was similar to that described by Horak[7] with an anode diameter of 1.3 cm, cathode diameter 0.7 cm, and an anode-cathode spacing 1–2 mm.

Prior to assembly, the cathode and anode parts were outgassed by heating in vacuum to a temperature of 1050°C or greater. The assembly of the tube, and the sealing-in and bake-out on exhaust at 450°C for 1 hour were similar to the usual oxide cathode procedures. The vacuum system used employed a three-stage oil diffusion pump and the vacuum was monitored with an ionization gauge. On cooling after bake-out the cathode was heated gradually to a temperature of about 1000°C allowing the pressure to go no higher than 5×10^{-6} mm Hg. The anode was also raised to between 900 and 1000°C by induction heating during this outgassing. After the pressure had been reduced to 5×10^{-7} mm, an anode potential of 70 volts was applied. The emission current was initially less than one microampere, but this current gradually increased and in some tubes increased up to emission current densities of 500 ma/sq cm at 1200°K.

The diodes were sealed off the system at a pressure of 5×10^{-8} mm or less and the tube was "gettered" using a Batalum type getter. Occasionally during processing it was noted that a black deposit formed on the tantalum anode, opposite the cathode, and could not be removed by heating. The presence of such a film appeared to poison the emission. The change to a platinum facing on the anode made it possible to remove the deposit by heating. It was noted in all diodes that the emission was stable even with the anode operating at red heat.

The zero field thermionic emission from the cathode was measured by the retarding potential method[7] for a series of different cathode temperatures. Assuming the zero field emission j_0 to vary with temperature according to the relation

$$j_0 = AT^2 e^{-e\Phi/kT}, \tag{1}$$

a plot of j_0/T^2 *vs* $1/T$ can be made, the slope of the straight line will give the value of the Richardson work function $e\Phi$, and the intercept gives the value of A.

[5] O. W. Richardson, Phil. Mag. **22**, 669 (1911).
[6] G. W. Fox and F. M. Bailey, Phys. Rev. **59**, 174 (1941).
¶ Nickel, 1001 pure Nickel obtained from International Nickel Company, New York. Platinum, obtained from Sigmund Cohn, New York.
[7] F Horak, J. Appl. Phys. **23**, 346 (1952).

The emission of the diode under pulse conditions was measured using a $3\mu s$ pulse at 60 pps with anode voltages between 60 and 600 volts. The current pulse was amplified (gain 2000) and current and voltage pulses were displayed on a synchroscope. Schottky plots were used to obtain the zero field emission for a series of temperatures and as in the case of the dc measurements, a Richardson plot could then be made. The apparent work function obtained in this way was about the same as that from dc measurements on the same cathode. The zero field current at 1000°K for pulse measurements was a factor of 3 greater than that for the dc measurements.

Table I is a compilation of the results of cathode materials including details of diode construction and indicating apparent work functions and A values. Column II indicates the method of preparation either from the hydrosulfide with external reduction to the sulfide, or by applying the hydrosulfide directly on the cathode and heating in vacuum on exhaust to form the sulfide, or from the sulfate by reduction in a hydrogen atmosphere. Column III lists the various cathodes tested with the details of anode and cathode face materials indicated in columns IV and V. Column VI shows the zero field emission densities (from dc measurements) at 1000°K in amp/sq cm, and the corresponding apparent work functions and A values are shown in columns VII and VIII. The apparent work functions from pulse measurements are indicated in column IX.

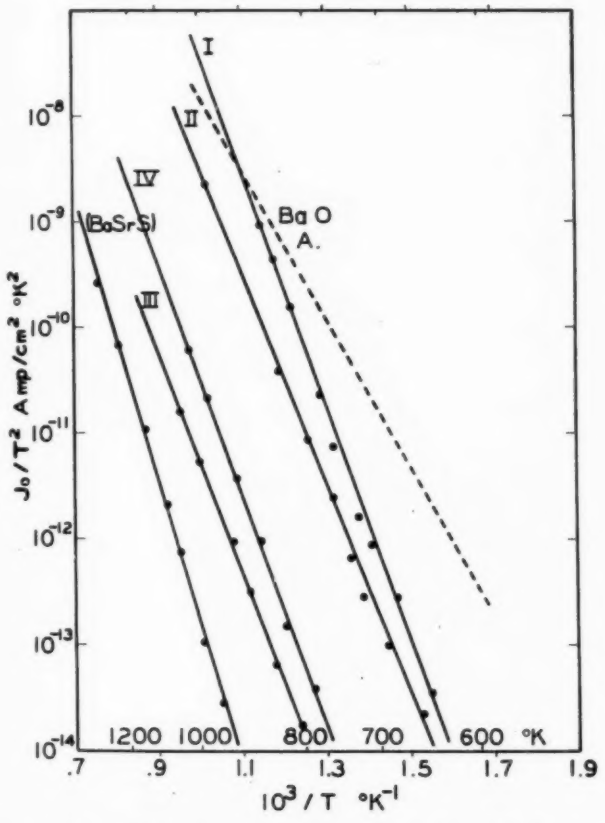


FIG. 1. Richardson plots of (BaSr)S, BaO, and BaS prepared from $Ba(SH)_2$.

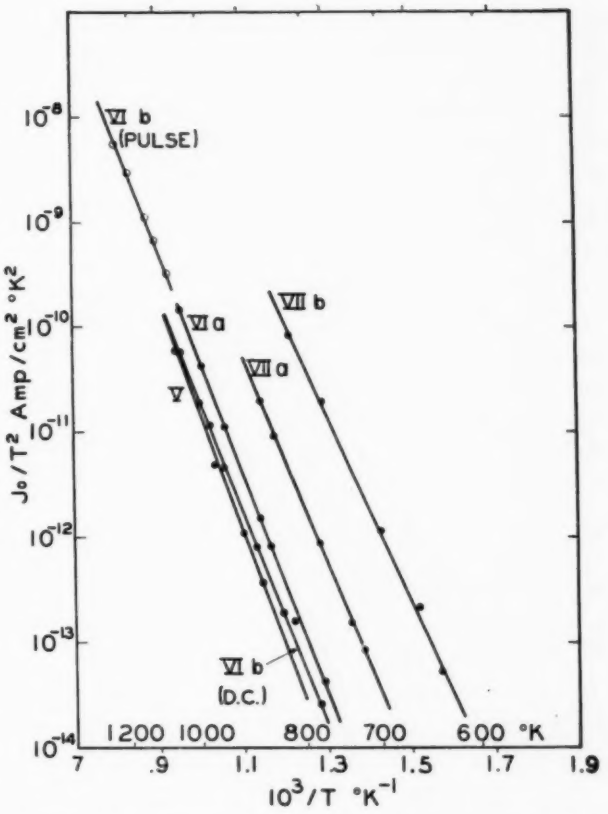


FIG. 2. Richardson plots of BaS prepared by reduction of $BaSO_4$.

Cathode I was made from barium sulfide material prepared initially by method (*a*). This particular coating material contained quite high impurity content of iron, tin, lead, chromium, and other metals. When materials of higher purity were prepared, the abnormally high A value was never repeated.

Cathode VI (*b*) was a cathode prepared from the pure barium sulfide of VI (*a*) with the addition of 300 ppm Fe. The iron was added to the barium sulfate in the form of ferrous chloride before reduction in hydrogen.**

Cathode VII (*b*) was a cathode prepared from the pure barium sulfide of VII (a) with the addition of 30 ppm Fe. In this case the iron was added to the barium sulfate in the form of ferrous sulfate before reduction in hydrogen.**

Included in Table I are values for a 50–50 molar mixture of (BaSr)S prepared by reduction of the equimolar mixture of the sulfate in hydrogen. Its apparent work function was 2.64 ev. No results were obtained for other Ba to Sr ratios. For comparison with the sulfide values emission data for three barium oxide cathodes[8] are shown.

** Colorimetric analyses for Fe after the reduction confirmed that this amount had been carried through.

[8] I. L. Sparks, Ph.D. thesis, University of Missouri (1951).

Figures 1 and 2 show Richardson plots for the various cathodes listed in Table I and enable comparison of the sulfide and oxide of barium for thermionic emission. It will be seen that the apparent work function of barium sulfide is 2.1 ev, but the activation possible varies between wide limits. Cathodes I, II, VII (*a*), and VII (*b*), all lie within the range of thermionic activity of the barium oxide cathodes at 1000°K.

(2) Electrical Conductivity

Because of its instability in air, very little reliable work has been done on the electrical conductivity of BaS. Fischer[9] in 1927 reported that BaS would not conduct electricity. His temperatures went no higher than 120°C, and he would not have been able to measure a specific conductivity less than 10^{-6} ohm^{-1} cm^{-1}. Gudden and Pohl[10] discussed photoconductivity in BaS, although no extensive work was undertaken.

The electrical conductivities of compressed pills of barium sulfide in high vacuum have been measured over a wide range of temperatures. The tube used for these measurements is shown in Fig. 3. The compressed pill is held between two electrodes supported from a circular glass press which itself is rigidly supported from a four-lead press sealed into the glass envelope. This assembly is held in the center of a tantalum cylinder (diameter 2.5 cm and 7.5 cm long) which is heated by means of radiofrequency power in the coil placed around the tube. A molybdenum-nickel†† thermocouple is used to monitor the temperature of the lower electrode, which is taken to represent the temperature of the pill. The electrodes were made of tantalum and the ends of the pills were coated with platinum either by evaporation in vacuum or by using a disk of 0.001-in. platinum foil between the pill and the tantalum. The tube was sealed on the vacuum system and baked out at 450°C before measurements were made.

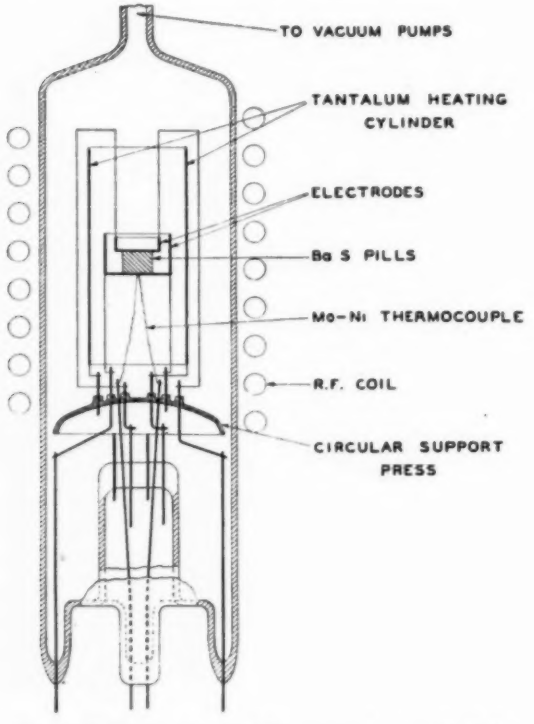


FIG. 3. Furnace tube used for measuring conductivity of BaS compressed pills in high vacuum.

The pills were 0.84-cm diameter and 0.5 cm thick. Their density was around 3.2 g/cm^3 or approximately 75 percent single crystal density. The conductivity was measured by applying a potential across the pill and measuring the current for different temperatures, care being taken to use a voltage well within the linear region of the voltage-current characteristic curve. The voltage applied was 1 volt or less. From a knowledge of the dimensions of the pill the specific conductivity σ could be calculated. Figure 4 includes plots of log σ *vs* $1/T$ for two pills in different states of activation. Curves 1 and 2 are such plots taken on a pure BaS pill (prepared from the same batch of material as Cathode VI (*a*)). The initial run with increasing temperature followed curve 1 (dots). An attempt was then made to activate this pill by passing 3 ma at 190 volts applied for several hours with the temperature at 1050°K. Only a small change in conductivity resulted (curve 2, triangles). Curves 1 and 2 have slopes which correspond to activation energies Q of 1.8 and 1.6 ev, respectively, where the conductivity is given by the equation

$$\sigma = \sigma_0 e^{-Q/kT}. \tag{2}$$

Curves 3 and 4 are plots of data taken on a BaS pill containing 300 ppm Fe impurity (from the same material as that used in Cathode VI (*b*)). Curve 3 (crosses) was taken after the bake-out on the vacuum system without any attempt at activation. The curve was reproducible with increasing or decreasing temperature, the slope of the intermediate temperature region (850-1000°K) corresponding to a Q value of 3.2 ev. This pill was activated by applying 5 to 20 volts for 5 hours at a temperature of 1050°K. The current passed at the maximum state of activity was 30 ma. Curve 4 was then taken showing activation by a factor of 100, and a slope corresponding to a value for Q of 1.5 ev.

The variation of electrical conductivity with temperature has also been measured for sprayed coatings of barium sulfide by placing two coated cathodes in spring contact in vacuum. This is similar to the method used extensively for barium oxide by Young.[11] The two cathodes with separate heaters are mounted on tungsten rod supports which keep them pressed together throughout the temperature range from room temperature to 1000°C. The cathodes had a platinum base on which the barium sulfide was sprayed and the temperature of

[9] P. Fischer, Z. Electrochem. **33**, 576 (1927).
[10] B. Gudden and R. Pohl, Z. Physik **16**, 44 (1923).

†† Calibration supplied by Dr. J. R. Young, private communication.
[11] J. R. Young, University of Missouri, Ph.D. thesis (1951).

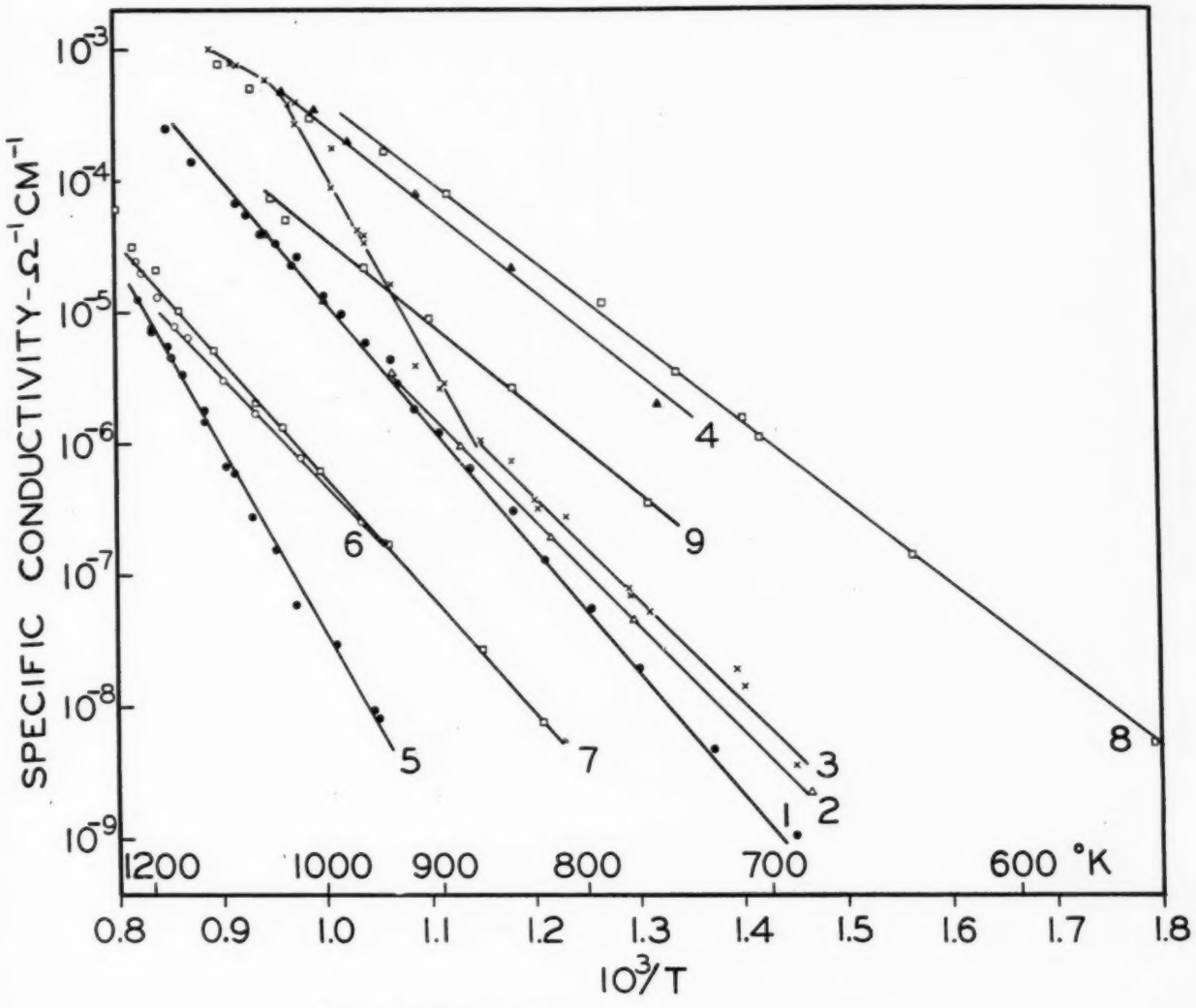


FIG. 4. Specific conductivity of BaS *vs* reciprocal temperatures.

Compressed pills

(1) BaS inactive $Q=1.8$ ev
(2) BaS active $Q=1.6$ ev
(3) BaS+300 ppm Fe inactive $\begin{cases} Q_1=3.2 \text{ ev} \\ Q_2=1.6 \text{ ev} \end{cases}$
(4) BaS+300 ppm Fe active $Q=1.5$ ev

Sprayed coatings

(5) BaS inactive $Q=2.8$ ev
(6) BaS active $Q=1.6$ ev
(7) BaS+30 ppm Fe inactive $Q=1.8$ ev
(8) (9) BaS+30 ppm Fe active $Q=1.2$ ev

each cathode was monitored by a Pt–Pt 10 percent Rh thermocouple. The thickness of material in this case was much smaller than for the pills, being about 0.4 mm, and the voltages used in the conductivity measurements were also much less than those necessary for the pills. The density of such sprayed coatings was about 1 g/cm³. The tubes were baked-out on the vacuum system at 450°C and the heaters outgassed by heating to a temperature somewhat higher than that used for conductivity measurements.

Figure 4, curve 5 shows a plot of $\log\sigma$ *vs* $1/T$ for pure BaS (the same material that was used for Cathode VII (*a*)) after the bake-out and outgassing procedure. The Q value for this curve is 2.8 ev. Curve 6 is a plot after the BaS had been activated by applying 4 volts for 5 hours at a temperature of 1200°K. The current drawn at maximum activation was 60 microamp and the Q value for this curve is 1.6 ev.

Plots of $\log\sigma$ *vs* $1/T$ for BaS with 30 ppm Fe impurity (the same material as that used in Cathode VII (*b*)) are also shown in Fig. 4. Curve 7 was taken after the bake-out and outgassing procedure without any attempt at activation. The Q value for this curve is 1.8 ev. Curve 8 was taken after activation at 1150°K by applying 2 volts for several hours. The maximum current drawn was 3 milliamp and the Q value for this curve is 1.2 ev.

Certain similarities and differences in the results obtained by the two methods of measuring conductivity should be noted:

(1) The activity of the unactivated specimens of BaS and Fe was of the same order as the activity of the activated pure BaS specimens.

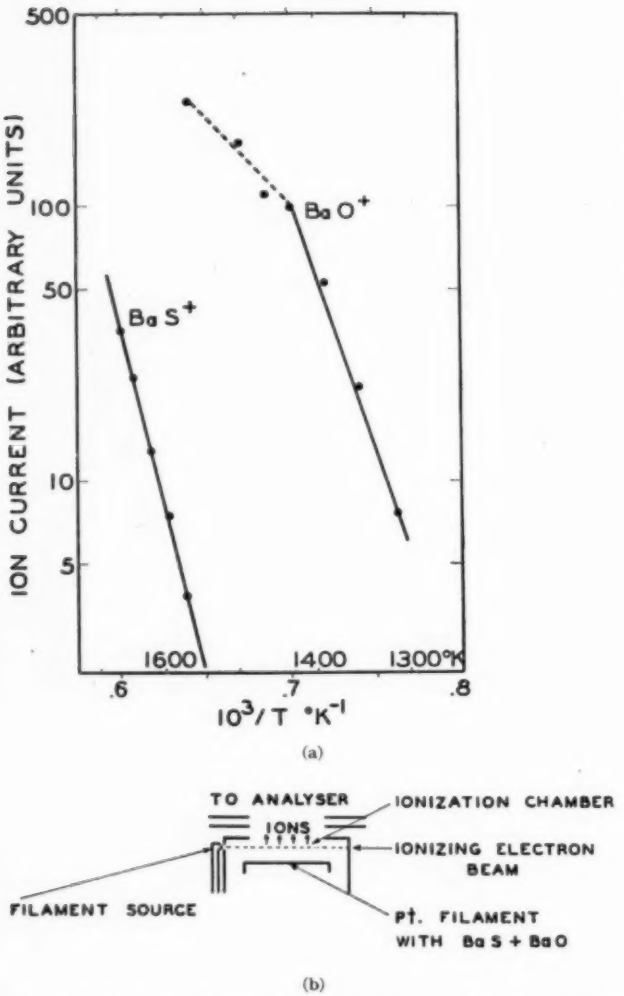


FIG. 5. (a) Log of ion currents of BaS^+ and BaO^+ *vs* reciprocal temperature, showing the relative vapor pressures of BaS and BaO. (b) Diagram of the mass spectrometer source.

(2) In the case of the pure BaS specimens, activation changed the activity by a factor less than 10.

(3) In the case of the BaS and Fe specimens, activation was much easier and resulted in increases by factors of 100–1000.

(4) The specific conductivity of each specimen was obtained from the geometry but no allowance was made for density. The density of the compressed pills was about 3 times that of the sprayed coatings. The results indicate a relation between specific conductivity and density. A similar effect has been noted with specimens of differing density of magnesium oxide.

(5) The compressed pill of BaS with Fe impurity before activation showed two different conductivity regions, whereas the sprayed cathodes before activation showed only one such region. Note that the Fe content of the compressed pills was much higher than that of the sprayed coatings.

(6) Activation produced a decrease in the activation energy of all specimens.

(3) Thermoelectric Power

The double cathode tube described in Sec. (*b*) was also used to measure thermoelectric emf's and their variation with temperature. As previously indicated, the temperature of each cathode could be separately controlled, hence a temperature difference could be maintained across the sample about a mean temperature. The thermoelectric emf was measured for a set of temperature differences and a plot of emf *vs* temperature difference was linear with the slope giving the thermal emf per degree, dE/dT, at the mean temperature. By this method, compensation could be made for any discrepancies in the two thermocouples due to asymmetry in position.

The emf's were measured using a Leeds and Northrup type K-2 potentiometer with a galvanometer of sensitivity 4×10^{-10} amp/mm to indicate balance. The use of this indicator restricted measurements to the temperature range in which the resistance of the coating was less than 10^9 ohms. The warmer electrode was always positive, indicating negative charge carriers.

The value of dE/dT was determined at several temperatures for the pure BaS in two states of activation and for the BaS+30 ppm Fe impurity in two states of activation. The value of dE/dT was found to be constant, and it was observed that the lower states of activity gave the higher values of dE/dT. Table II is a summary of the thermoelectric data. When the determination of dE/dT on the activated BaS and Fe was begun, the material had an activity corresponding to curve 8, Fig. 4; however, when the conductivity was redetermined after the thermoelectric power data were taken, the conductivity had decreased (curve 9, Fig. 4).

TABLE II. Thermoelectric data for BaS and BaS+30 ppm Fe for unactivated and activated specimens.

Material	State of activity	Temperature range	dE/dT
Pure BaS	Curve 5, Fig. 4	700–850°C	2.7 ±0.1 mv per degree
Pure BaS	Curve 6, Fig. 4	770–1000°C	2.3 ±0.2 mv per degree
BaS+30 ppm Fe	Curve 7, Fig. 4	475–850°C	1.4 ±0.1 mv per degree
BaS+30 ppm Fe	Curve 8, 9, Fig. 4	475–850°C	1.1 ±0.2 mv per degree

The following points are suggested by the experimental results:

(1) The thermionic emission and conductivity results appear to be capable of interpretation by a semiconductor picture of barium sulfide as an N-type semiconductor.

(2) The sign of the thermoelectric effect indicates negatively charged carriers in agreement with the above concept, and also its value decreases with increase in activity. However, the results show the thermoelectric power to be constant with temperature whereas on the semiconductor picture, the temperature dependence of the conductivity and thermoelectric power should be similar. A similar constant thermoelectric

power with temperature was found by Young[11] for BaO in the low temperature region. As yet there is no explanation of these results.

(4) Evaporation Studies

A 50–50 mixture of BaS–BaO was applied to the platinum strip source of a Nier mass spectrometer in a manner similar to that used by Aldrich.[12] The source of the spectrometer is shown diagrammatically in Fig. 5(a). As this platinum strip is heated by passing current through it, material will be evaporated into the ionizing chamber in the path of the electron beam directed at right angles to the axis of the spectrometer. The electron beam will ionize the evaporated material producing BaS^+ and BaO^+ ions and also Ba^+ ions by dissociation and ionization. These positive ions drift up into the slit system of the spectrometer, are accelerated, analyzed according to their e/m ratio, and then collected. If the assumption is made that the ionization efficiencies for BaS and BaO by 75 volt ionizing electrons are the same, then a comparison of BaS^+ and BaO^+ ion currents at any temperature is the same as a comparison of the vapor pressures of the two materials at that temperature.

Figure 5(b) is a plot of the ion currents *vs* $1/T$, and it will be seen that the BaS^+ ion currents are smaller than the BaO^+ by about two orders of magnitude, indicating a correspondingly lower vapor pressure for the sulfide. It can also be seen that in order to obtain a similar amount of evaporation from BaS as one gets from BaO operated at 1400°K, the sulfide must be operated at 1600°K. The slope of the plots in Fig. 5(b) is a measure of the heat of sublimation and the values from these slopes are

82 kcal/mole for BaO
104 kcal/mole for BaS.

Such measurements indicate that the mass spectrometer can be useful in the comparison of vapor pressures and in the estimation of heats of sublimation of refractory materials where more standard methods are inapplicable. It is based on the two assumptions that the probability of ionization by the ionizing electrons for the comparison materials is the same and that the ion currents produced are proportional to the vapor pressures of the materials. It would seem useful to explore such assumptions extensively on other materials for which the heats of sublimation are accurately known.

[12] L. T. Aldrich, J. Appl. Phys. **22**, 1168 (1951).

CONCLUSIONS

This study has established a Richardson work function for barium sulfide of 2.1 ev with a thermionic emission at 1000°K comparable with that from pure barium oxide. The variation of conductivity with temperature has been studied as well as the thermoelectric power. An n-type semiconductor model is capable of explaining the results apart from the constancy of the thermoelectric power with temperature. The vapor pressure of BaS is found to be much less than that of BaO with a value for the heat of sublimation of 104 kcal/mole.

The authors are pleased to acknowledge the guidance and encouragement of Professor A. S. Eisenstein and the assistance of their colleagues. One of us (Walter Grattidge) expresses appreciation for the hospitality extended him.

APPENDIX A

Spectrographic analyses are expressed in parts per million in Table III. The samples are designated by number of the cathode on which data were taken.

TABLE III

	I	II[b]	VIa[c]	$BaCO_3$[d]
Fe[a]	30	≪30	1	3
Al	>10	10	10	*t*
Mg	>20	30	20	50
Mn	*t*	*t*	...	0.5
Si	*p*	<3	2	*t*
Sn	20	...	...	...
Pb	>10	3	1	0.5
Na	10	...	...	10
Cu	5	<5	1	0.5
Ca	*p*	*p*	*p*	5
Sr	>30	30	<100	30

"*t*" indicates trace present.
"*p*" indicates that element was present but no estimation was made of its concentration.

[a] The Fe in samples VI*a*, VI*b*, and VII*b* were analyzed colorimetrically and contained 1 ppm, 300 ppm, and 30 ppm, respectively.
[b] This analysis is typical of all material prepared by the hydrosulfide method.
[c] This analysis is typical of all material made by the reduction of $BaSO_4$, and would include the material used on Cathodes V, VI*a*, and VII*a* and, with the exception of Fe, Cathodes VI*b* and VII*b*.
[d] Material used by Sparks, reference 8.

Scattering of Electromagnetic Energy in a Randomly Inhomogeneous Atmosphere

HAROLD STARAS
National Bureau of Standards, Washington, D. C.
(Received April 4, 1952)

First-order perturbation theory is used to evaluate the scattered power at a receiver resulting from random inhomogeneities in the propagating medium. The integral expression for this scattered power is equivalent to the expressions as used by Pekeris and Booker and Gordon. However, it is shown that it is not appropriate to use a space correlation function of refractive index as defined by the above authors. Instead, this paper defines a time correlation function of refractive index which permits formal evaluation of the time average scattered power. It is also shown that, whereas a space correlation function of refractive index is not amenable to direct experimental evaluation, the time correlation function as defined in this paper is directly measurable. Finally, it is shown that for so-called small-scale turbulence the average scattered power does not depend appreciably on any particular model of atmospheric turbulence, whereas for large-scale turbulence the frequency and scattering angle dependence of the scattered energy is affected greatly by the particular time correlation function chosen or by the assumption that scattering is the result of randomly located spheres.

I. THE HERTZ POTENTIAL FOR THE SCATTERED RADIATION

CONSIDER a nonconducting medium with dielectric constant ϵ which may be written

$$\epsilon(\mathbf{r}, t)=\epsilon_0+s\epsilon_1(\mathbf{r}, t) \quad s\ll 1,$$

where s is a small positive number and ϵ_0 is a constant representing the time average of ϵ. From Maxwell's equations it can be shown that the Hertz vector $\pi(\mathbf{r}, t)$ satisfies the equation

$$\nabla^2\pi+\frac{s\epsilon_1}{\epsilon_0}\nabla(\nabla\cdot\pi)-(\epsilon_0+s\epsilon_1)\mu\frac{\partial^2\pi}{\partial t^2}=-\frac{1}{\epsilon_0}\int\mathbf{J}dt. \quad (1)$$

For a homogeneous medium (i.e., $\epsilon_1=0$), Eq. 1 reduces to the well-known wave equation. With $\epsilon_1\neq 0$, Eq. (1) may be solved by perturbation theory where it is assumed that π may be expanded in a power series of s, i.e.,

$$\pi=\pi_0+s\pi_1+s^2\pi_2+\cdots,$$
$$\mathbf{E}=\mathbf{E}_0+s\mathbf{E}_1+s^2\mathbf{E}_2+\cdots,$$

where $\mathbf{E}_n=\nabla(\nabla\cdot\pi_n)-\epsilon_0\mu(\partial^2\pi_n/\partial t^2)$. Upon substitution into Eq. (1) and equating powers of s, we obtain the following two equations for π_0 and π_1:

$$\nabla^2\pi_0-\epsilon_0\mu\frac{\partial^2\pi_0}{\partial t^2}=-\frac{1}{\epsilon_0}\int\mathbf{J}dt, \quad (2a)$$

$$\nabla^2\pi_1-\epsilon_0\mu\frac{\partial^2\pi_1}{\partial t^2}=-\frac{\epsilon_1}{\epsilon_0}\mathbf{E}_0. \quad (2b)$$

Equation (2a) yields the solution of the wave equation for a homogeneous atmosphere, while Eq. (2b) is the inhomogeneous wave equation and has the solution

$$\pi_1(\mathbf{r}, t)=\frac{1}{4\pi\epsilon_0}\int_v\frac{\epsilon_1}{R'}\left(\mathbf{r}', t-\frac{R'}{c}\right)\mathbf{E}_0\left(\mathbf{r}', t-\frac{R'}{c}\right)dv', \quad (3)$$

where c is the velocity of light in a medium with capacitivity ϵ_0 (not free space) and $R'=|\mathbf{r}'-\mathbf{r}|$. Since $\mathbf{E}_0$ is the primary incident field, we see that the Hertz potential for the scattered radiation (to a first approximation) results from the elementary dipoles induced by the primary field.

The integral in Eq. (3) is over all of space. Let us divide our whole space into small macroscopic volumes in each of which the incident field is constant in magnitude and direction. π_1 may then be written

$$\pi_1=\frac{1}{4\pi\epsilon_0}\sum_i\int_{v_i}\frac{\epsilon_1}{R_i}(\mathbf{r}_i', t-R_i'/c)\mathbf{E}_{0i}(\mathbf{r}_i', t-R_i'/c)dv_i'. \quad (4)$$

From Eq. (4) and the relations connecting $\mathbf{E}$ and $\mathbf{H}$ with π, it can be shown that, neglecting the near fields,

$$\mathbf{E}_1=\sum_i\mathbf{E}_{1i}=-k^2\sum_i\mathbf{a}_i\times(\mathbf{a}_i\times\pi_{1i}),$$
$$\mathbf{H}_1=\sum_i\mathbf{H}_{1i}=k^2Y_0\sum_i\mathbf{a}_i\times\pi_{1i},$$

where $\mathbf{a}_i$ is a unit vector in the direction $\mathbf{r}-\mathbf{r}_i'$ and Y_0 is the characteristic admittance of a medium with capacitivity ϵ_0. Thus the scattered power density at the receiver during one cycle (ϵ_1 being considered constant over one cycle of the carrier) is given by $\mathbf{P}_1(\mathbf{r}, t)=\frac{1}{2}Re\mathbf{E}_1\times\mathbf{H}_1^*$, which is the real part of $\frac{1}{2}Y_0(k^2/4\pi)^2$ multiplied by

$$-\sum_{i,j}\int_{v_i}\int_{v_j}\frac{\epsilon_1(\mathbf{r}_i', t-R_i'/c)\epsilon_1(\mathbf{r}_j'', t-R_j''/c)\left[\mathbf{a}_i\times\left\{\mathbf{a}_i\times\mathbf{E}_{0i}\left(\mathbf{r}_i', t-\frac{R_i'}{C}\right)\right\}\right]\times\left[\mathbf{a}_j\times\mathbf{E}_{0j}\left(\mathbf{r}_j'', t-\frac{R_j''}{C}\right)\right]}{\epsilon_0\quad\epsilon_0\quad R_i'\quad R_j''}dv_i'dv_j''. \quad (5)$$

It might be well to indicate at this point that each term of the double sum in Eq. (5) is a random function since it involves integrals of random functions. Therefore $\mathbf{P}_1(\mathbf{r}, t)$ is a random function of position and time

The simplest statistical property of $\mathbf{P}_1(\mathbf{r}, t)$ to evaluate is its time average which is defined by

$$\langle \mathbf{P}_1 \rangle_{Av} = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \mathbf{P}_1(\mathbf{r}, t)dt$$

$$= \frac{Y_0}{2} \sum_i |\mathbf{E}_{0i}|^2 \left(\frac{k^2 \sin X_i}{4\pi R_i'}\right)^2 \mathbf{a}_i \int_{v_i}\int_{v_i} N(\mathbf{r}'-\mathbf{r}'')$$

$$\times \exp\{-jk[(R_0'+R')-(R_0''+R'')]\}dv'dv'', \quad (6)$$

where X_i is the angle between $\mathbf{a}_i$ and $\mathbf{E}_{0i}$. In Eq. (6) we have introduced the time correlation function of ϵ_1/ϵ_0, defined by

$$N(\mathbf{r}'-\mathbf{r}'')^{\dagger} = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \frac{\epsilon_1(\mathbf{r}', t)}{\epsilon_0} \frac{\epsilon_1(\mathbf{r}'', t)}{\epsilon_0} dt$$

$$\cong \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \frac{\epsilon_1(\mathbf{r}', t-R'/C)}{\epsilon_0} \frac{\epsilon_1(\mathbf{r}'', t-R''/c)}{\epsilon_0} dt.$$

The second approximate equality holds if ϵ_1 does not change measurably in a time of length $(R'-R'')/c$. In addition, it was assumed that $\mathbf{E}_{0i}(\mathbf{r}', t)$ may be written as $|\mathbf{E}_{0i}| \exp[j(\omega t - kR_0)]$, R_0 being the distance from the transmitter to the point of integration in v_i. Naturally, in all of the cross-product terms in Eq. (5) the time average of $\epsilon_1(\mathbf{r}_i', t)\epsilon_1(\mathbf{r}_j'', t)$ is zero except in the immediate neighborhood of the surface common to v_i and v_j. Clearly, this surface effect is negligible compared to the volume integrations in Eq. (6).

II. EVALUATION OF THE AVERAGE SCATTERED POWER

Since $\frac{1}{2}Y_0|E_{0i}|^2 = P_{0i}$ is the power incident on v_i, (6) may be rewritten

$$\langle \mathbf{P}_1 \rangle_{Av} = \sum_i \langle \mathbf{P}_{1i} \rangle_{Av} = Re \sum_i P_{0i} \mathbf{a}_i \left(\frac{k^2 \sin X_i}{4\pi R_i}\right)^2 I_i, \quad (7)$$

where

$$I_i = \int_v \int_{v_i} N(\mathbf{r}'-\mathbf{r}'')$$

$$\times \exp\{-jk[(R_0'-R_0'')+(R'-R'')]\}dv'dv''.$$

To evaluate I_i, imagine that the volume v_i is sufficiently far from the transmitter so that it may be assumed that a plane wave traveling in the $+z$ direction is incident on v_i. To this order of approximation $R_0'-R_0''=z'-z''$. With the origin of coordinates centered in v_i and the receiver at a distant point $x_0y_0z_0$

$$R' = [(x'-x_0)^2+(y'-y_0)^2+(z'-z_0)^2]^{\frac{1}{2}},$$

$$\cong d\left(1-\frac{x'x_0+y'y_0+z'z_0}{d^2}\right),$$

where $d=(x_0^2+y_0^2+z_0^2)^{\frac{1}{2}}$. Letting $\alpha=x_0/d$, $\beta=y_0/d$, $\gamma=z_0/d$,

$$R' \cong d-(\alpha x'+\beta y'+\gamma z').$$

A similar expression holds for R''. Therefore,

$$R'-R'' = -[\alpha(x'-x'')+\beta(y'-y'')+\gamma(z'-z'')].$$

Since the vector $\mathbf{r}'-\mathbf{r}''$ has components $x'-x''$, $y'-y''$, $z'-z''$, I_i may be written

$$I_i = \int_{-x_1}^{x_2} \cdots \int_{-x_1}^{x_2} N(x'-x'', y'-y'', z'-z'')$$

$$\times \exp\{jk[\alpha(x'-x'')+\beta(y'-y'')$$

$$+(\gamma-1)(z'-z'')]\}dx'\cdots dz'', \quad (8)$$

where, for simplicity, it is assumed that the volume of integration may be taken as a rectangular parallelepiped. If we leave the double-primed coordinates unchanged and introduce $\xi=x'-x''$, $\eta=y'-y''$, $\zeta=z'-z''$, Eq. (8) becomes

$$I_i = \int_{-x_1}^{x_2}\int_{-y_1}^{y_2}\int_{-x_1}^{x_2}\int_{-(x_1+x'')}^{x_2-x''}\int_{-(y_1+y'')}^{y_2-y''}\int_{-(x_1+x'')}^{x_2-x''}$$

$$\times N(\xi, \eta, \zeta)e^{jk[\alpha\xi+\beta\eta+(\gamma-1)\zeta]}d\xi\cdots dz''. \quad (9)$$

In Eq. (11) let us integrate over ξ and x'' first, i.e.,

$$J = \int_{-x_1}^{x_2}\int_{-(x_1+x'')}^{x_2-x''} N(\xi, \eta, \zeta)e^{jk\alpha\xi}d\xi dx''.$$

The inner integral over ξ is a function of x'' since the limits depend on x''. Let us call the inner integral $F(x'')$. Integrating by parts

$$J = x''F(x'')\Big|_{-x_1}^{x_2} - \int_{-x_1}^{x_2} x'' \frac{dF}{dx'}(x'')dx'',$$

which, upon evaluation, yields

$$J = \int_{-(x_1+x_2)}^{x_1+x_2} (x_1+x_2-|\xi|)N(\xi, \eta, \zeta)e^{jk\alpha\xi}d\xi.$$

Thus a double integral has been reduced to a single integral. It is clear then that the sixfold integral in (9) can be reduced to the triple integral

$$I_i = v_i \int_{-(x_1+x_2)}^{x_1+x_2}\int_{-(y_1+y_2)}^{y_1+y_2}\int_{-(z_1+z_2)}^{z_1+z_2}$$

$$\times\left(1-\frac{|\xi|}{x_1+x_2}\right)\left(1-\frac{|\eta|}{y_1+y_2}\right)\left(1-\frac{|\zeta|}{z_1+z_2}\right)$$

$$\times N(\xi, \eta, \zeta)e^{jk[\alpha\xi+\beta\eta+(\gamma-1)\zeta]}d\xi d\eta d\zeta. \quad (10)$$

Equation (10) is an exact expression for the double-

† Note that the correlation function is explicitly assumed to be a function of the vector distance between the two points at which the time correlation of ϵ_1 is evaluated.

volume integral in (6). However, it can be further simplified under some conditions. If, for $|\xi|$, $|\eta|$ or $|\zeta| \geq L$, $N(\xi, \eta, \zeta)$ is essentially zero and if all three dimensions of the scattering volume, namely x_1+x_2, y_1+y_2, z_1+z_2 are much greater than L, then Eq. (10) may be rewritten with negligible error as

$$I_i = v_i \int_{-\infty}^{\infty}\int\int N(\xi, \eta, \zeta) e^{jk[\alpha\xi+\beta\eta+(\gamma-1)\zeta]} d\xi d\eta d\zeta. \quad (11)$$

If, now, N is assumed to be an isotropic function, i.e., a function of $r=(\xi^2+\eta^2+\zeta^2)^{\frac{1}{2}}$, then Eq. (11) can be rewritten (see Appendix II) in the form

$$I_i = \frac{4\pi v_i}{\Gamma} \int_0^{\infty} N(r) \sin(\Gamma r) r dr, \quad (12)$$

where $\Gamma = 2k \sin(\Theta/2)$, $k=2\pi/\lambda$, Θ = angle between the direction of the energy incident on v_i and the direction of the receiver from v_i. This is formally identical with the result of Booker and Gordon[1] as given by the real part of the equation one line above their Eq. (2*s*). It should be recalled, however, that N has been defined differently from ρ as given by Pekeris[2] and Booker and Gordon[1] and it is only in the sense as defined in this paper that Eq. (12) is correct. If $N(r)$ is set equal to $N(0)e^{-r/l}$, where $N(0)=\langle(\epsilon_1/\epsilon_0)^2\rangle_{Av}$ is the time average of $(\epsilon_1/\epsilon_0)^2$, then I_i as used in this paper is the same as the I used by Booker and Gordon.

III. SMALL-SCALE TURBULENCE

If $N(r)$ is essentially zero for $r \geq L$ and $L \ll \lambda/2\pi$, then it is said that the scale of turbulence is small. For this case, the major contribution to the integral in Eq. (12) comes when $\Gamma r = (4\pi r/\lambda)\sin\Theta/2$ is a small quantity. The sine may then be replaced by its argument and Eq. (12) becomes

$$I_i = 4\pi v_i \int_0^{\infty} N(r) r^2 dr.$$

From the above it is clear that for small-scale turbulence, the integral, I_i, is independent of scattering angle and frequency but its amplitude would certainly depend on the specific form of $N(r)$.

IV. LARGE-SCALE TURBULENCE

Several simple functions will be used as possible functional forms for $N(r)$, and it will be shown that the wavelength and scattering angle dependence of I_i is quite different for the different functional forms of $N(r)$. The following forms will be chosen:

$$N(r) = N(0)\exp(-r/l), \quad (13a)$$

$$N(r) = N(0)\exp[-(r/l)^2], \quad (13b)$$

$$N(r) = N(0)/[1+(r/l)^2]^2, \quad (13c)$$

$$N(r) = N(0)[1-(r/l)]^2 \quad r \leq l \quad (13d)$$

$$= 0 \quad r > l,$$

where l is an appropriate scale factor. For the foregoing cases Eq. (12) becomes

(a) $$I_i = \frac{8\pi l^3 v_i N(0)}{[1+(\Gamma l)^2]^2},$$

(b) $$I_i = \pi^{\frac{1}{2}}\pi l^3 v_i N(0) \exp\left\{-\left(\frac{\Gamma l}{2}\right)^2\right\},$$

(c) $$I_i = \pi^2 l^3 v_i N(0)\exp\{-\Gamma l\}, \text{ and}$$

(d) $$I_i = 4\pi l^3 v_i \frac{N(0)}{(\Gamma l)^4}\left[4 - \frac{6}{(\Gamma l)}\sin(\Gamma l) + 2\cos(\Gamma l)\right].$$

Each of the foregoing cases is quite different one from the other when Γl varies from zero to fairly large values (e.g., 20). It would thus appear that before reliance can be placed on this approach to the problem of scattering by random inhomogeneities, an experimentally reliable form for N must be obtained. A graphical comparison of the above results are shown in Fig. 1. Since the correlation function given by Eq. (13a) has already been suggested by previous authors,[1-3] Fig. 1 presents the intensity of the scattered radiation for the different

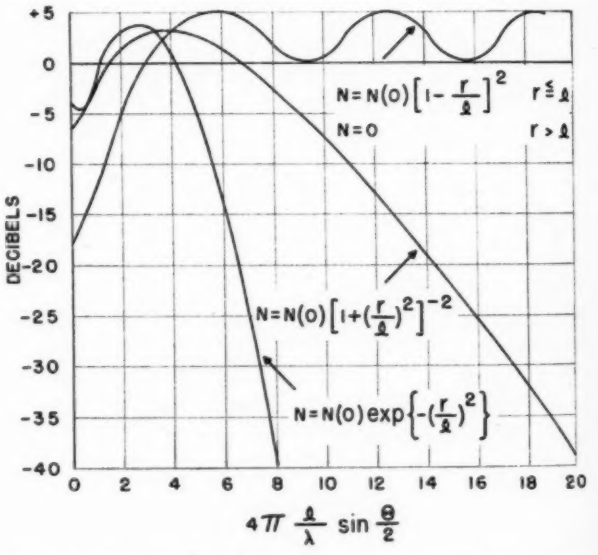


FIG. 1. The dependence of the scattered radiation on the correlation function. The scattered radiation is measured in decibels above that obtained when the correlation function is $N(0)\exp\{-r/l\}$.

[1] H. G. Booker and W. E. Gordon, Proc. Inst. Radio Engrs. **38**, 401 (1950).
[2] C. L. Pekeris, Phys. Rev. **71**, 268 (1947).
[3] George Birnbaum, Phys. Rev. **82**, 110 (1951).

correlation functions expressed in db relative to that obtained for Eq. (13a).

V. SCATTERING CROSS SECTION

At this point, it might be convenient to define a scattering cross section per unit volume for the volume v_i. Using Eq. (12), Eq. (7) can be rewritten

$$\langle \mathbf{P}_i \rangle_{Av} = \sum_i \frac{\mathbf{P}_{0i}}{R_i^2} \mathbf{a}_i v_i \sigma_i, \tag{14}$$

where

$$\sigma_i = \frac{(k^2 \sin X_i)^2}{4\pi\Gamma} \int_0^\infty N(r) \sin(\Gamma r) r dr.$$

Inspection of Eq. (7) shows that σ_i is the scattering cross section per unit volume, per unit incident power, per unit solid angle. It thus has the same dimensions as the scattering cross section defined by Booker and Gordon.[1] It is defined here, however, in terms of the time average scattered power.

VI. SCATTERED BY SPHERES

In scattering by spheres, where the cross-sectional profile of refractive index has the same form as $N(r)$, the scattering cross section of each sphere would be given by

$$\sigma_s = \left(\frac{k^2 \sin X_i}{4\pi}\right)^2 \left| \int_v N(r) \exp[-jk(R_0+R)] dv \right|^2$$

$$= \left(\frac{k^2 \sin X_i}{4\pi}\right)^2 \left[\frac{4\pi}{\Gamma} \int_0^\infty N(r) \sin(\Gamma r) r dr\right]^2.$$

The average scattering cross section for a volume containing n spheres per unit volume randomly located would then be n times as great (simply adding the powers in a unit volume):

$$\sigma_i = \left(\frac{k^2 \sin X_i^2}{4\pi}\right)^2 n \left[\frac{4\pi}{\Gamma} \int_0^\infty N(r) \sin(\Gamma r) r dr\right]^2. \tag{15}$$

For small-scale turbulence, Eq. (15) would differ from Eq. (14) at most by a multiplicative factor since the integrals in both of these equations become independent of frequency and scattering angle. The two expressions for σ_i could be made identical merely by adjusting n relative to $N(0)$. However, for large scale turbulence, they would in general differ in their dependence on frequency and scattering angle.

To obtain the average total power scattered, one must perform the summation in Eq. (7) or approximate the summation by an integral. When taking the gain of the transmitting and receiving antennas into account, the average available scattered power to the receiver $\langle P_s \rangle_{Av}$ may be written as

$$\langle P_s \rangle_{Av} = P_0 \lambda^2 \int \frac{g_t g_r}{(4\pi R_0)^2 R^2} dv,$$

where g_t and g_r are the gains of the transmitting and receiving antennas and are functions of direction from the transmitter and receiver, P_0 is the radiated power and $\lambda^2 g_r/4\pi$ is the absorbing area of the receiving antenna.

VII. CONCLUDING COMMENT

In the previous mathematical development two assumptions (among others) have been made in order to obtain a fairly simple expression for the scattering cross section of a portion of space. One assumption is that space may be broken up into macroscopic volume elements sufficiently small so that the magnitude and direction of the incident electric field may be considered constant. The other assumption is that these same macroscopic volumes have linear dimensions much larger than the scale of turbulence. Great care must, therefore, be used in applying the results of this paper to propagation over an earth. In propagation typical of FM and TV broadcasting with high transmitting antennas, the lobe structure due to the ground-reflected ray may be such that it is impossible to find volume elements satisfying both of the above conditions. In this case, as well as in the case when the time correlation is not isotropic, the exact expression for I_i as given in Eq. (10), rather than the simpler expression found in Eq. (12), may prove useful. However, it is important not to divide space into volume elements which are too small since then the time average of the cross-product terms in Eq. (5) will not necessarily be negligible.

Grateful acknowledgement is extended to K. A. Norton, J. Feinstein, and C. H. Page of the National Bureau of Standards and J. B. Keller of New York University Mathematics Department for many helpful suggestions and discussions.

APPENDIX I. EVALUATION OF THE DOUBLE INTEGRAL

$$I = \int_v \int_v \frac{\epsilon_1}{\epsilon_0}(\mathbf{r}, t) \frac{\epsilon_1}{\epsilon_0}(\mathbf{r}', t) \times \exp\{-jk[(R_0 - R_0') + (R - R')]\} dv dv'.$$

It is the purpose of this appendix to point out this writer's objections to the mathematical evaluation of the above integral as presented in references 1 and 2. In reference 2, it is stated: "To evaluate I we make the following change. At present the coordinates of dv at P and of dv' at P' are both measured from an origin 0, and either of the two volume integrations could be carried out first. Let us now regard dv' as the inner integration and let the coordinate of P be measured relative to P' instead of 0." The foregoing integral may then be rewritten in the manner of Booker and Gordon:

$$I = \int \left[\int_v \frac{(\epsilon_1)}{\epsilon_0} \frac{(\epsilon_1)'}{\epsilon_0} dv' \right] \times \exp[-jk\{(R_0 - R_0') + (R - R')\}] dv;$$

but

$$\rho=\frac{1}{v\left\langle\left(\frac{\epsilon_1}{\epsilon_0}\right)^2\right\rangle_{Av}}\int_v\left(\frac{\epsilon_1}{\epsilon_0}\right)\left(\frac{\epsilon_1}{\epsilon_0}\right)'dv';$$

hence

$$I=v\left\langle\left(\frac{\epsilon_1}{\epsilon_0}\right)^2\right\rangle_{Av}\int_v\rho \times\exp[-jk\{(R_0-R_0')+(R-R')\}]dv.$$

The above mathematical development is open to the following two objections. In the first case, when integrating over dv' to obtain ρ, the vector distance between $\epsilon_1(\mathbf{r})/\epsilon_0$ at P and $\epsilon_1'(\mathbf{r}')/\epsilon_0$ at P' must be maintained constant throughout the entire volume. Careful inspection of the above double integral will show that it is impossible to keep the vector distance between P and P' constant and at the same time keep $[(R_0-R_0')+(R-R')]$ constant. This error is negligible, however, in integrating over a small volume sufficiently far from both transmitter and receiver. In the second case, and this is more serious, since P is being measured from P', the limits on the variables of integration of dv depend on P' and it is not permissible to integrate over dv' first. There is, of course, a more complicated way of manipulating the double volume integral, I, which would permit integration over dv' first while keeping the coordinate differences, $P-P'$, constant. But in this case there would result, in effect, eight double-volume integrals instead of just one. Furthermore, the volume of integration as well as the separation between the two points would depend on $P-P'$. As a matter of fact for some values of $P-P'$, the volume of integration for dv' becomes vanishingly small. Integrating a random function over a very small volume will obviously yield a random result rather than something as well defined as a correlation function, as is implied in the derivation presented in references 1 and 2.

The reader can convince himself of the above by considering the simpler case of the following double integral,

$$\int_a^b\int_a^b\epsilon(x)\epsilon(x')\exp[jk(x-x')]dxdx',$$

and introducing $\xi=x-x'$, i.e., measuring P from P'.

It might be appropriate at this point to indicate that the space correlation function for the refractive index of the atmosphere (as defined in references 1 and 2) is not amenable to *direct* experimental evaluation, while the time correlation function as defined in this paper is already being measured at the Bureau of Standards by G. Birnbaum and his colleagues. To make this point clear, consider the space correlation function, ρ, defined above. To evaluate that integral experimentally, it would be necessary to perform many volume integrations (one for each separation between the points P and P') over fairly large volumes of space in a period of time so short that ϵ_1 does not vary in time. This is not possible with today's experimental techniques. However, the time correlation function as defined in this paper can be measured directly by setting up two refractometers[4-5] at two points in space and correlating the refractive index as it varies in time at these two points.

APPENDIX II. EVALUATION OF THE INTEGRAL

$$\int_{-\infty}^{\infty}\int\int N(\xi,\eta,\zeta)\exp\{jk[\alpha\xi+\beta\eta+(\gamma-1)\zeta]\}d\xi d\eta d\zeta.$$

It is here assumed that $N(\xi,\eta,\zeta)=N(r)$, where $r^2=\xi^2+\eta^2+\zeta^2$. For this case it is most convenient to introduce spherical coordinates. Now, $\alpha\xi+\beta\eta+(\gamma-1)\zeta=\mathbf{r}\cdot\mathbf{A}=|\mathbf{A}|r\cos\theta$, where $\mathbf{A}$ is a fixed vector with components α, β, $\gamma-1$. It should be remembered that α, β, γ are the direction cosines of the receiving location from the scattering volume. In particular $\gamma=\cos\Theta$, where Θ is the angle between the direction of propagation and the direction of the receiver from the scattering volume. Therefore,

$$|\mathbf{A}|=[\alpha^2+\beta^2+(\gamma-1)^2]^{\frac{1}{2}}=[2(1-\cos\Theta)]^{\frac{1}{2}}=2\sin(\Theta/2).$$

Choosing the polar axis along $\mathbf{A}$, the integral in question may be written

$$2\pi\int_0^\infty N(r)r^2\left[\int_0^\pi\exp(j\Gamma r\cos\theta)\sin\theta d\theta\right]dr,$$

where $\Gamma=2k\sin(\Theta/2)$. The inner integral becomes $2\sin\Gamma r/\Gamma r$. The remaining integral is, therefore,

$$\frac{4\pi}{\Gamma}\int_0^\infty N(r)\sin(\Gamma r)rdr.$$

[4] G. Birnbaum, Rev. Sci. Instr. **21**, 169 (1950).
[5] C. M. Crain, Rev. Sci. Instr. **21**, 456 (1950).

Microchronometric Schlieren Study of Gaseous Expansion from an Electric Spark*

H. L. Olsen, R. B. Edmonson, and E. L. Gayhart
Applied Physics Laboratory, The Johns Hopkins University, Silver Spring, Maryland
(Received May 14, 1952)

Gaseous phenomena induced by short-duration sparks at atmospheric pressure have been photographed during the first microseconds of their existence by a flash-schlieren technique. The growth of the heated gas domain (kernel) expanding from the spark discharge has been observed for air and argon. The kernel continues to expand rapidly, but subsonically, after the spark current ceases and after the shock wave separates. It develops further into a torus about the spark electrodes and eventually decays. The development of flame from a heated gas kernel in a flammable mixture is illustrated. Energy relationships for the expanding argon kernel are discussed.

INTRODUCTION

EXTENSIVE work has been done on the spark ignition of gaseous combustibles. Treatments of spark ignition have been published by Lewis, von Elbe, Guest, and Blanc[1–3] and by Fenn,[4] which represent the latest attacks on the problem. The thermal ignition theories developed therein are shown to be in fair accord with experimental values of minimum spark-ignition energies, quenching distances, and lean flammability limits. The summation of these investigations has resulted in an over-all understanding of the principles involved in the process. However, direct observation and analysis of the ignition and propagation of flames from sparks during the early phases of the process has received little attention. Motion picture techniques yield a limited number of frames with relatively long exposure times and thus are not suitable for a detailed study of the process as it exists during the first millisecond. This paper reports experimental observations on the heated gas domain (kernel) which is induced by an electric spark and from which a flame is observed to propagate in a flammable mixture. The heated kernel expanding from the path of the electrical discharge is studied photographically during the first microseconds of its existence as well as during its later development. The experimental techniques utilized make possible a detailed examination of the phenomena, particularly in the early phases, which has not been done before. The nature of the electrical discharge through the gas *is not* in question here.

EXPERIMENTAL TECHNIQUES

Apparatus

The observations and measurements reported derive from flash schlieren photographs of a gaseous spark discharge and the associated gaseous phenomena at atmospheric pressure. As shown in Fig. 1 the ignition spark is made to pass between electrodes mounted in a free vertical jet of gas in the schlieren field. A precision interval timer (varitimer) provides accurate control of the time between the firing of the ignition spark and the illuminating Libessart spark so that any phase of the development of the phenomena following the ignition spark from coincidence to one thousand microseconds may be photographed.[5] Electromagnetic radiation from the ignition spark is picked up by the varitimer which subsequently fires a Libessart spark controlled by a hydrogen thyratron in series with the gap. The effective exposure time for the Libessart gap is about one-quarter microsecond with the result that the rapidly changing phenomena are resolved. The effect of the finite light duration, however, is observable as a slight blurring of the pattern for the most rapidly changing phenomena (shock waves). The dioptric schlieren system, shown diagrammatically in Fig. 1, produces bilateral symmetry in the schlieren image by use of a bilateral light-source slit and a fine wire as a schlieren stop. This system produces excellent definition but does not show the sense of the density gradients. The choice of focal lengths results in a magnification of the schlieren field of 3.4:1 on the camera screen.

The gases under observation were supplied from commercial cylinders through flow-control apparatus to the free vertical jet mentioned above. The adjustable spark electrodes were so arranged that the spark and subsequent phenomena occurred wholly within the gas

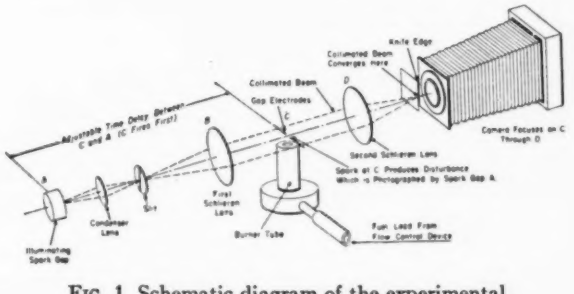


FIG. 1. Schematic diagram of the experimental schlieren arrangement.

* Work supported by Bureau of Ordnance, U. S. Navy, NOrd Contract 7386.

[1] Blanc, Guest, von Elbe, and Lewis, J. Chem. Phys. **15**, 798 (1947).

[2] B. Lewis, and G. von Elbe, J. Chem. Phys. **15**, 803 (1947).

[3] Blanc, Guest, von Elbe, and Lewis, *Third Symposium on Combustion and Flame and Explosion Phenomena, Madison, Wisconsin* (The Williams and Wilkins Company, Baltimore, Maryland, 1949), p. 363.

[4] John B. Fenn, Ind. Eng. Chem. **43**, 12 (1951).

[5] Edmonson, Gayhart, and Olsen, J. Soc. of Phot. Engrs. **3**, 3 (1952).

jet. Thus, within limits of present observation, the spark occurs in an atmosphere of uniform composition and pressure.

Characteristics of the Electric Spark

The ignition spark was produced by the discharge of a condenser through an "overvolted" gap of pointed electrodes. Control of the spark was obtained by triggering a hydrogen thyratron (Type 5C22), in series with the spark gap, with a pulse produced by either a manual push button or a varitimer. The arrangement was similar to that used for the Libessart gap.

The sparks involved in the work of references 1, 2, 3, and 4 are simple condenser discharges with minimized external circuit losses. The energy dissipated at the gap under such conditions is a maximum and can be compared to the potential energy stored within the condenser.[6] For our experiments, the variation and unpredictibility of the uncontrolled spark-gap breakdown is a disadvantage, and for this reason the aforementioned control system has been used. The thyratron provides precision control of the spark-gap breakdown but introduces the disadvantage of energy dissipation in the thyratron. Circuit losses, other than that introduced by the thyratron, have been minimized.

For simplicity of interpretation of the results, a short duration spark is desirable, and it thus becomes necessary to investigate the effect of the hydrogen thyratron upon the time duration of the spark current. The current pulse has been determined utilizing the voltage drop across a small series resistor in the cathode return of the thyratron. This voltage is displayed as a function of time on the screen of a synchroscope, and the resulting current-time trace photographed. A drawing, traced directly from the oscillogram, is shown in Fig. 2. The time from grid trigger to anode current is 0.55 microsecond. The value quoted by Heins[7] for this thyratron is 0.50 to 0.90 microsecond, depending upon the stiffness of the trigger pulse. The length of the current pulse proper is 0.45 microsecond. Oscillation or "ringing" caused by slight cable impedance mismatch continues for another 0.3 microsecond. Thus the time duration of the spark is effectively less than 0.45 microsecond, since the effect of the cathode resistor used in this test is to lengthen the time of current flow. These measurements of current pulse duration were made utilizing a *GE* 0.01 microfarad Pyranol condenser. For the vacuum condensers used in part of this work, the current pulse time certainly does not exceed the aforementioned value.

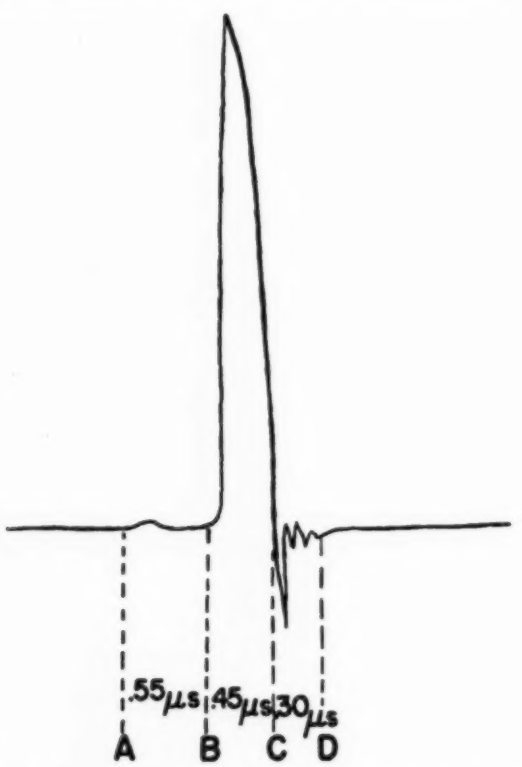


FIG. 2. Current-time trace for ignition spark. A. Thyratron grid firing pulse. B, C. Anode current pulse. C, D. Cable oscillation. This illustration has been traced from an oscillogram.

RESULTS AND DISCUSSION

The development of the phenomena associated with an ignition-spark discharge in air during its very early phases is shown in Fig. 3. The spark was obtained from a bank of vacuum condensers (2175 mmf) charged to 8000 volts. The first picture of this series shows the spark gap, without an ignition spark, silhouetted by the schlieren light source. The subsequent series of pictures spaced at quarter-microsecond intervals shows the systematic development of the gas phenomena associated with the spark up until and after the separation of a shock wave from the hot gas kernel left in the wake of the discharge. Each photograph shows a separate and indpendent spark. The high degree of reproducibility of the phenomenon makes the series of photographs equivalent to a motion picture of a single spark at a rate of four million frames per second. In order to obtain these pictures it was necessary to use two varitimers tripped simultaneously by radiation pick-up from a third and independent spark discharge. One varitimer controls the ignition spark and the other the Libessart gap. The time from ignition spark to schlieren photograph is thus the difference in the settings of the two varitimers. With this arrangement it is possible to flash the schlieren light source even before the ignition spark. No attempt has yet been made to correlate the beginning of the current trace of the spark with the appearance of the schlieren pattern; hence, for the time quoted in this article, the picture preceding the first appearance of the schlieren pattern has been taken

[6] Roth, Guest, von Elbe, and Lewis, J. of Chem. Phys. **19**, 12 (1951).

[7] Harold Heins, *Electronics Engineering Manual* (McGraw-Hill, Inc., New York), Vol. V, pp. 246–252.

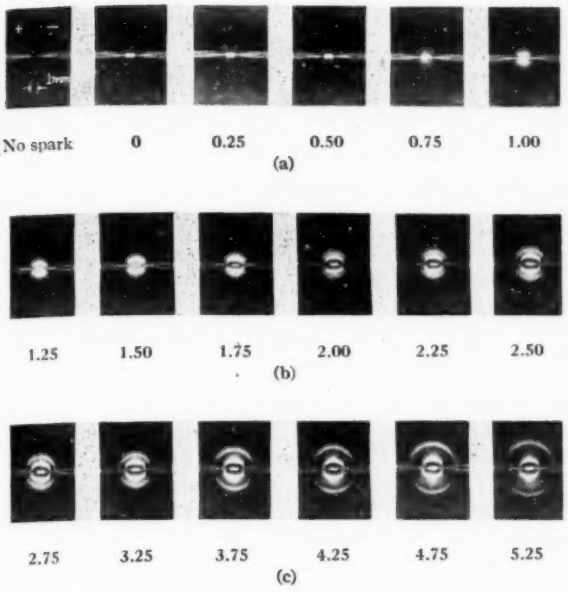


FIG. 3. The formation of the heated kernel and shock wave from an electric spark in air. Each photograph is of a different spark. View perpendicular to axis of electrodes. Time of photograph given in microseconds.

arbitrarily as a reference zero time. In this series of pictures the direct light from the spark, which is recorded on the film as the bright line between the electrodes, is not to be confused with schlieren patterns. This same direct-light exposure appears in each picture of the series since the camera shutter is open for one-fiftieth of a second for each spark-synchronized exposure. Thus any light emitted by the spark or the hot gases will record on film independently of the schlieren system if its intensity is above the threshold for the film. The second picture ($t=0$ microsecond) of the series shows the direct spark light but no schlieren pattern, thus indicating that the schlieren exposure was made before refractive index gradients were introduced by the spark. The next picture ($t=0.25$ microsecond) shows a faint pattern surrounding the direct exposure. This pattern persists with increased intensity to approximately one microsecond. Thereafter a rapid change sets in and a steady development of the inner core ensues. By 2.00 microseconds it is apparent that a shock wave has separated from the residual hot gases and is traveling into the surrounding atmosphere. The shock wave as well as the hot gas kernel remaining behind the shock shows a well-defined boundary which is amenable to accurate measurement. The width of the boundaries, as seen on the film, is partially determined by the effective duration of the light source.

The nature of the faint pattern which precedes the rapid development stage has not been investigated. Since the pattern does not appear if the schlieren light is flashed early enough, and since it does rotate with the schlieren stop, its source is not direct radiation from the spark region. Any direct radiation proceeding from the region of this pattern would mark the film without regard to the timing, as does the direct spark light between the electrodes. This observation is corroborated by schlieren pictures taken later when the kernel has been swept clear of the electrodes by the gas stream.

The rapid development of the kernel for several microseconds is interpreted as deriving from the delivery of spark energy to the gas for about one-half microsecond after spark breakdown occurs. The duration of current flow, being less than half a microsecond, encompasses only two successive photographs. The remainder of the development thus occurs in the absence of current flow. After the separation of the shock wave, the residual heated gas kernel continues to expand subsonically. The further development of the kernel is shown in three successive stages in Fig. 4. In these pictures two views, taken perpendicular to and parallel to the axis of the electrodes, show the complete geometry of the kernel at each phase of its development. At 15 microseconds, the kernel is an ellipsoid and the concentric shock wave is still in the field of view. By 112 microseconds the kernel has developed into a torus. The kernel thus possesses the characteristics of a closed vortex, as does a smoke ring. The striated pattern appearing in the side view within the torus was observed to reverse when the polarity of the electrodes was reversed, thus indicating that the sense of the vortical motion is determined by the direction of current flow. The mass motion required for the development of the torus may (possibly) arise from impulsive forces exerted by the electric field on the ions produced by the spark discharge. The toroidal shape has been consistently obtained for all gases examined including flammable mixtures. The toroidal development is also observed with and without the thyratron in series with the spark and for condensers of widely differing dielectrics (e.g., vacuum condensers and oil condensers).

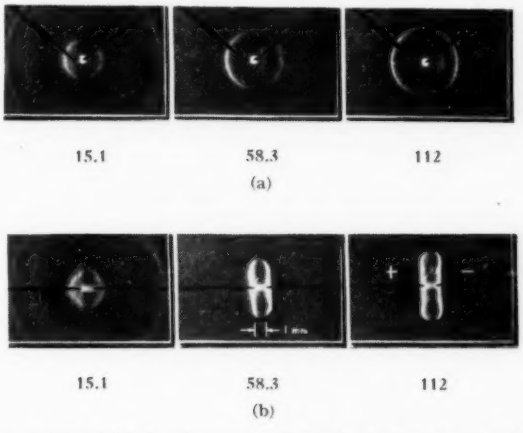


FIG. 4. Toroidal development of the heated air kernel from a spark. Spark gap with series hydrogen thyratron, 7000 volts, 0.01 mf Pyranol condenser. Each photograph of a different spark. Top view, along electrode axis; bottom view, perpendicular to electrode axis. Time in microseconds.

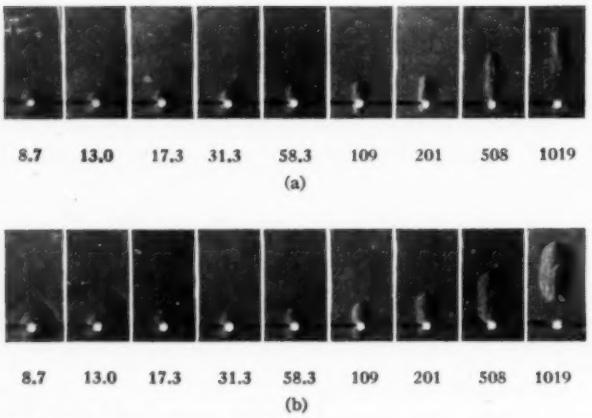


FIG. 5. Air and flame kernels from an electric spark. Top, the development of a heated kernel from an electric spark in a stream of air; bottom, the development of a flame kernel from an electric spark in a stoichiometric propane-air stream. Ignition spark identical in the two series. Time given in microseconds.

Ignition of Propane-Air Flame

When the air jet is replaced by a jet of flammable mixture, ignition occurs when the spark energy exceeds a minimum ignition energy for the particular flammable mixture. Such an ignition in a propane-air mixture is shown in Fig. 5. This figure compares the growth of an air kernel with that of the incipient flame for identical sparks and delay times. The steady growth of the flame as contrasted to the eventual decay of the air kernel is apparent. The development of a torus is also evident in the intermediate stages of the kernel growth in both flame and air. Unlike the other photographs of this report, the flame pictures were taken with a schlieren system equipped with a conventional single knife edge. The resulting pictures are not bilaterally symmetric, and the relative sense of the horizontal density gradients is apparent. These pictures show the appearance of the spark electrodes after the kernel has been swept away by the air jet and thus corroborate the statements made earlier on the subject of light emission from the kernel.

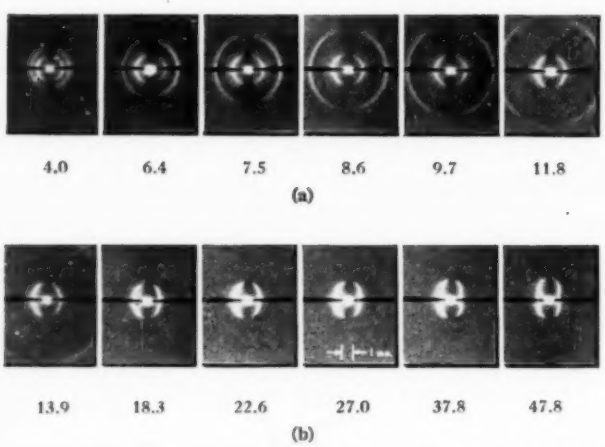


FIG. 6. Development of heated argon kernel from a spark. Spark gap with series thyratron. 7000 volts, 0.01 mf. Each snapshot of a different spark. View perpendicular to the electrode axis. Time given in microseconds.

Kernel Behavior in Argon

A detailed history of the growth of a heated kernel resulting from a spark in argon during its expansion phase after shock-wave separation is shown in Fig. 6. Argon was used in order to avoid dissociation effects and other energy processes associated with polyatomic molecules. When the actual ignition process is considered these effects must be taken into account. The kernel was viewed only perpendicular to the spark axis, in which aspect the kernel outline appears as an ellipse. The major and minor axes were measured, and the volumes of the ellipsoids were computed as those generated by revolving the ellipses about the axis of the spark electrodes. The resulting kernel volume is displayed as a function of time in Fig. 7. This curve shows that the expansion of the kernel proceeds rapidly but at a decreasing rate for the first ten microseconds or so, at which time it is apparently slightly over-expanded. After a slight decrease in volume it proceeds to expand slowly again, and this slow expansion ultimately leads to the development of the aforementioned torus and eventually kernel decay. The rapid expansion of the kernel after the shock separation, however, is subsonic. This general pattern of behavior has been observed to hold for several gases (N_2, A, air, and propane-air mixtures).

The continued rapid expansion of the kernel after the spark current has ceased flowing and the shock wave has separated requires explanation. The expansion is too rapid for interpretation in terms of diffusion of the heated gas into the surrounding atmosphere, and this explanation is also incompatible with the sharply defined kernel boundary. The expansion would appear to be associated with a changing pressure in the gases within and surrounding the kernel. There can be no appreciable pressure differential acro.s the kernel boundary since that boundary is observed to move subsonically with respect to the undisturbed atmosphere. Therefore the entire mass of gas behind the shock wave must be at pressures above atmospheric, and this general elevated pressure must drop as the shock wave progresses. This viewpoint is substantiated by measurements made on the shock wave. The shock-wave velocity has been determined as a function of time from measured values of the shock-wave diameters from Fig. 6. This velocity is supersonic at 4.00 microseconds (just after separation of kernel and shock wave) with a Mach number of 1.5 and decreases as the shock wave propagates away from the kernel. The velocity becomes indistinguishable from sonic velocity (Mach number unity) in about 16 microseconds, which is in accord with the kernel expansion. From the value of the Mach number at 4.00 microseconds, the static pressure ratio across the shock is 2.56. This value for the ratio is to be

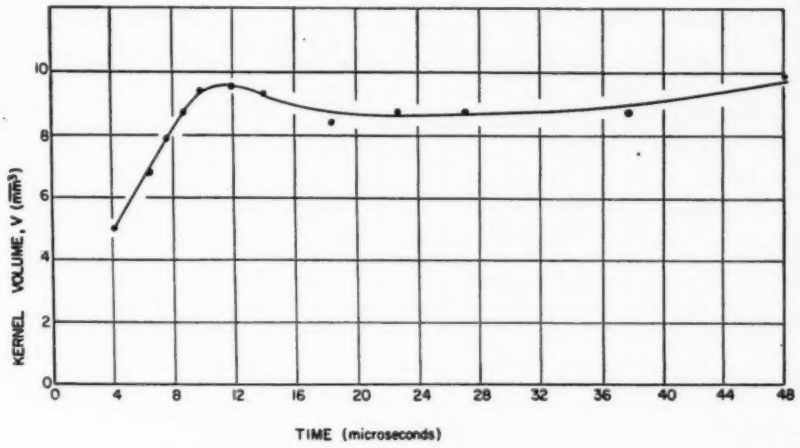


FIG. 7. Volume of argon kernel as a function of time. Data from measurements of photographs of Fig. 6.

compared with the pressure ratio $[(V_f/V_s)^\gamma=2.52]$ computed from the adiabatic law of expansion utilizing kernel volumes measured at 4.00 microseconds and at pressure equilibration. This agreement is somewhat better than is expected from the accuracy of the measurements. From this result it is concluded that the kernel expansion is approximately represented by the adiabatic law from the time of shock-wave separation until the time of pressure equilibration.

The condenser energy $1/2\ CV^2$ is utilized in the following manner:

$$1/2\ CV^2-E_x=E_L+E_s, \tag{1}$$

in which E_x is the loss in the circuit external to the spark gap, E_s is the fraction of the condenser energy existing as internal energy of the kernel at the time of shock wave separation, and E_L the loss from the kernel up to this time. Because of the obvious lack of thermal equilibrium and the uncertainty as to the physical conditions prevailing before shock-wave separation, no attempt has been made to evaluate the quantity E_L. In reality, thermal equilibrium probably does not exist within the kernel at shock separation; however, the subsequent kernel expansion is surprisingly well represented by the adiabatic law. On the basis of the foregoing measurements, the assumption is made that the adiabatic law holds, i.e., that one may ascribe a temperature to the kernel gases and that there is, therefore, no change in mass M_s of the kernel during this time. With these restrictions an analytical expression can be obtained for the internal energy of the kernel at the time of shock-wave separation together with an expression for the change in internal energy due to the subsequent expansion. Let the state variables be V_s, P_s, and T_s at the time of shock-wave separation. Define V_0 as the volume of M_s corresponding to the initial gas conditions T_0 and P_0. Refer to the $P-V$ diagram of Fig. 8. From the time of spark discharge to shock-wave separation the mass of gas has been transformed from the condition $A(V_0, P_0, T_0)$ to $D(V_s, P_s, T_s)$ along an unknown path. Since the internal energy change from A to D is independent of the path, one may evaluate this quantity along the path ABD:

$$E_s=\Delta U|_{AB}+\Delta U|_{BD}, \tag{2}$$

where

$$\Delta U|_{AB}=\rho_0 V_0 c_p(T_x-T_0)-P_0(V_s-V_0), \tag{3}$$

$$\Delta U|_{BD}=\rho_0 V_0 c_v(T_s-T_x), \tag{4}$$

in which ρ_0 is the initial gas density, c_p and c_v are the heat capacities, and T_x is a temperature between T_0 and T_s. Noting that along BD

$$T_x=\frac{T_s P_{,0}}{P_s} \tag{5}$$

and using

$$\frac{T_s}{T_0}=\frac{P_s V_s}{P_0 V_0}=\frac{V_f{}^\gamma}{V_0 V_s{}^{\gamma-1}} \quad \text{and} \quad \rho_0 c_v T_0=\frac{P_0}{\gamma-1}, \tag{6}$$

in which the subscript f refers to the time of pressure

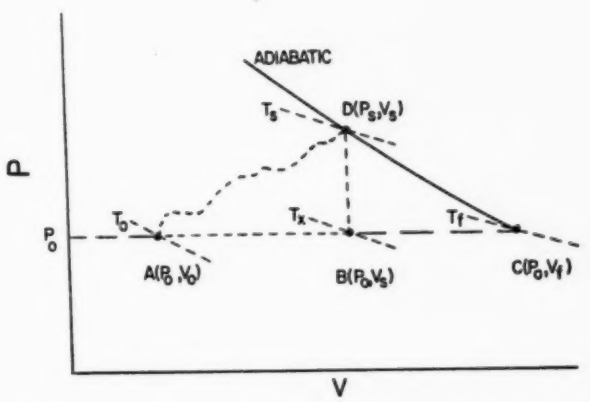


FIG. 8. $P-V$ diagram for expansion of argon kernel. The unknown path from A to D is indicated by the irregular dotted lines. D is the point just after shock-kernel separation, and C is the point of pressure equilibration.

equilibration at C, then

$$E_s=\frac{P_0}{\gamma-1}\left[\frac{V_f^{\gamma}}{V_s^{\gamma-1}}-V_0\right]. \tag{7}$$

A similar calculation yields the internal energy of the kernel at pressure equilibration as

$$E_f=\frac{P_0}{\gamma-1}[V_f-V_0]. \tag{8}$$

Since V_0 is small compared to both V_s and V_f, one obtains a good approximation for the internal energy from Eqs. (7) and (8) if we neglect V_0. The energy expended from D to C is given by

$$E_s-E_f=\int_{V_s}^{V_f} PdV=\frac{P_0}{\gamma-1}\left(\frac{V_f^{\gamma}}{V_s^{\gamma-1}}-V_f\right). \tag{9}$$

The fractional loss of energy because of the expansion is

$$\frac{E_s-E_f}{E_s}=\frac{(V_f/V_s)^{\gamma-1}-1}{(V_f/V_s)^{\gamma-1}-V_0/V_f}. \tag{10}$$

Therefore, if V_0/V_f is neglected in (10), one obtains a *minimum* fractional energy loss given by

$$\frac{E_s-E_f}{E_s}=\frac{(V_f/V_s)^{\gamma-1}-1}{(V_f/V_s)^{\gamma-1}}. \tag{11}$$

Thus it follows that the change in internal energy during the adiabatic expansion can be calculated together with good approximations to E_s and E_f. This method, applied to circuits where E_x is minimized or to those where $1/2\ CV^2-E_x$ is measured, thus makes available a method for evaluating E_L. Utilizing the measured value of $V_f/V_s\doteq 1.7$ from Fig. 7, (11) becomes

$$\frac{E_s-E_f}{E_s}\doteq 0.3.$$

This calculation indicates that the internal energy of the kernel varies with time and that part (as shown for argon, a considerable part) of the energy transferred from the spark is distributed throughout the surrounding infinite atmosphere as the work required to expand the kernel. That part of the energy remaining in the kernel is available for ignition purposes.

The authors wish to thank Dr. F. T. McClure, Dr. W. H. Avery, and Dr. W. G. Berl for their interest in this work and for valuable suggestions made during discussion of the paper.

JOURNAL OF APPLIED PHYSICS VOLUME 23, NUMBER 10 OCTOBER, 1952

Effect of Internal Strains on Linear Expansion, X-Ray Lattice Constant, and Density of Crystals

P. H. MILLER, JR., AND B. R. RUSSELL
Department of Physics, University of Pennsylvania, Philadelphia, Pennsylvania
(Received May 21, 1952)

General formulas are derived for the shift in x-ray diffraction maxima due to small distortions in simple cubic crystals. The special case of radial distortions due to lattice defects is considered in detail; it is found that the relative change in lattice constant computed from the shift in diffraction maxima is of the order of magnitude of the defect concentration and larger by as much as 100 percent than the linear expansion when the concentration of defects is of the order of one percent or less. A comparison is made of lattice constant, linear expansion, and density expected in crystals containing Frenkel defects, Schottky defects, or F centers.

INTRODUCTION

THE principal assumptions made in calculating the shift in x-ray diffraction maxima due to small radial distortions are given in the first section. A more general treatment given in Appendix 1 shows that the symmetry arguments upon which the simplified treatment is based are justified. In the second section the effect of a random distribution of defects, each of which produces a small radial distortion, is estimated in various ways; by averaging the values obtained by methods which are thought to overcorrect the results for symmetric crystallites with those which undercorrect, a formula is obtained for the calculated change in lattice constant accurate to about 10 percent. The corresponding increase in linear dimensions is derived in Appendix 2; in Appendix 3 the effect of Schottky and Frenkel defects is compared, and in Appendix 4 the change in density due to F centers is computed. The significance of these results is brought out in the third section where experimental results are discussed and experiments which may yield some new information about lattice defects in ionic crystals are proposed.

I. SYMMETRIC CRYSTALLITES

The distortion in a simple cubic crystallite of identical atoms having a lattice vacancy (or interstitial atom) at the center may be assumed to be such that all atoms in the crystallite are displaced radially from their normal positions, as shown in exaggerated form in Fig. 1. The change in x-ray lattice constants produced by radial strain of this sort is readily calculated in simple cases. Let the normal positions of the atoms be denoted by the vector set

$$\mathbf{r}=L_1\mathbf{a}_1+L_2\mathbf{a}_2+L_3\mathbf{a}_3,$$

having as their reciprocal system the vector set

$$\mathbf{B}=h_1\mathbf{b}_1+h_2\mathbf{b}_2+h_3\mathbf{b}_3,$$

where $L_1L_2L_3$ and $h_1h_2h_3$ are integers and $\mathbf{a}_j\cdot\mathbf{b}_k=\delta_{jk}$. The vector describing the displacement of each atom is assumed to be radial and hence may be written in the form

$$\Delta\mathbf{r}=\mathbf{r}f(r).$$

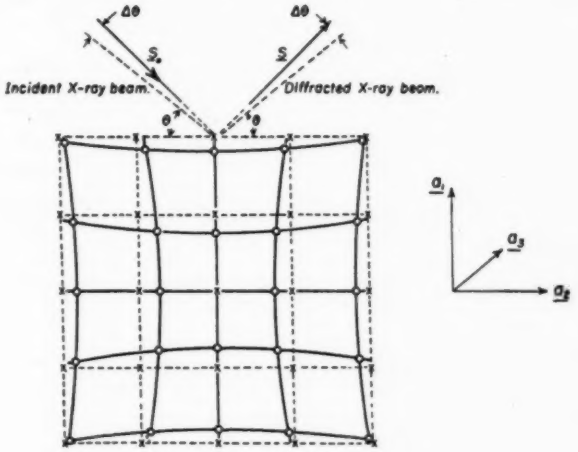
The positions occupied by the atoms in the distorted crystallite are then given by the vectors

$$\mathbf{R}=\mathbf{r}+\Delta\mathbf{r}=\mathbf{r}[1+f(r)].$$

The phase of the x-rays coherently scattered by an atom at vector position $\mathbf{R}$ relative to those scattered by an atom at the origin is given by

$$\delta(\mathbf{R})=\frac{2\pi}{\lambda}(\mathbf{S}-\mathbf{S}_0)\cdot\mathbf{R},$$

where $\mathbf{S}_0$ is a unit vector in the direction of the incident beam and $\mathbf{S}$ is a unit vector in the direction of scattering under consideration. The scattered waves all have the same phase in the case of the idealized crystal for which $\mathbf{R}=\mathbf{r}$ and the vector $\mathbf{S}-\mathbf{S}_0/\lambda$ coincides with a reciprocal lattice vector, as the phase differences are then multiples of 2π for every scattered wave; the determination of the values of $\mathbf{S}-\mathbf{S}_0/\lambda$ which satisfy this condition, hence producing strong "reflection,"



X — Positions of Atoms in undistorted Crystal : $\underline{r}=L_1\underline{a}_1+L_2\underline{a}_2+L_3\underline{a}_3$

O — Positions of Atoms due to radial distortion : $\underline{R}=\underline{r}(1+f(r))$

θ = Bragg angle for undistorted Crystal, $n\lambda=2d\sin\theta$,

when diffraction maximum occurs at $\theta+\Delta\theta$, calculated

lattice constant changes : $\frac{\Delta d}{d}=-\cot\theta\Delta\theta$

FIG. 1. Illustration of radial distortion in a simple cubic crystal.

thus locate points in reciprocal space from which the values of the **a** vectors are then computed. The distortion of the lattice produces a shift in the directions in which strong reflections occur, and hence the calculated values of the **b** and **a** vectors are affected. In order to investigate the effect quantitatively, let us first consider a symmetric crystallite having a lattice defect at the origin. As a further simplifying assumption, suppose that the crystal is so mounted that $\mathbf{S}-\mathbf{S}_0/\lambda$ coincides in direction with $\mathbf{a}_1$ and hence also with $\mathbf{b}_1$ in this special case (the usual Bragg spectrometer mount, for instance). In the undistorted crystal, the strong reflections would then occur at angles of incidence such that $\mathbf{S}-\mathbf{S}_0/\lambda=h_1\mathbf{b}_1+0+0$ or to reflection from crystal planes having Miller indices (h_100). It can be seen from symmetry arguments that the effect of radial distortion will be to produce a strong reflection at a slightly different angle of incidence but that $\mathbf{S}-\mathbf{S}_0$ will still coincide in direction with $\mathbf{a}_1$ and $\mathbf{b}_1$. From the shift in the diffraction maximum a new value of $\mathbf{b}_1$ would then be computed; the change can be expressed in the form

$$h_1\mathbf{b}_1'=(h_1+\Delta h_1)\mathbf{b}_1,$$

since the two vectors are parallel, and it is readily shown that the relative change in the direct lattice vector $\mathbf{a}_1$ which we shall denote by $\Delta d/d$ is given by

$$\Delta d/d=-\Delta h_1/h_1.$$

The problem then resolves itself into finding that value of Δh_1 which yields a maximum in the intensity of the wave diffracted from the distorted crystallite. Assuming the waves scattered from each atom to be identical in amplitude the expression for the relative amplitude of the total wave is

$$\begin{aligned}E(\mathbf{S},\mathbf{S}_0,\lambda)&=\sum_{L_1L_2L_3}\exp\frac{2\pi i}{\lambda}(\mathbf{S}-\mathbf{S}_0)\cdot\mathbf{R}(L_1L_2L_3)\\&=\sum_{L_1L_2L_3}\exp 2\pi i(h_1+\Delta h_1)\mathbf{b}_1\cdot[\mathbf{r}(1+f(r))]\\&=\sum_{L_1L_2L_3}\exp 2\pi i(h_1+\Delta h_1)L_1(1+f(r))\\&=\sum_{L_1L_2L_3}\exp 2\pi iL_1[h_1f(r)+\Delta h_1(1+f(r))],\end{aligned}$$

since $\exp 2\pi iL_1h_1=1$. If $f(r)$ and Δh_1 are sufficiently small that the products $L_1\Delta h_1$ and $L_1h_1f(r)$ are much less than unity, then to a first approximation,

$$E\doteq\sum_{L_1L_2L_3}\left(1+x+\frac{x^2}{2}\right)=\sum_{L_1L_2L_3}\left(1+\frac{x^2}{2}\right)$$

where

$$\begin{aligned}x(\Delta h_1)&=2\pi iL_1[h_1f(r)+\Delta h_1(1+f(r))]\\&\doteq 2\pi iL_1[h_1f(r)+\Delta h_1].\end{aligned}$$

The term in x vanishes by symmetry. The diffraction maximum occurs at

$$\frac{\partial E}{\partial(\Delta h_1)}=0=\frac{\partial}{\partial(\Delta h_1)}\{\sum_{L_1L_2L_3}x^2\}=\sum_{L_1L_2L_3}2x\frac{\partial x}{\partial(\Delta h_1)}.$$

Hence,

$$\sum_{L_1L_2L_3}L_1^2(h_1f(r)+\Delta h_1)=0,$$

$$h_1\sum_{L_1L_2L_3}L_1^2f(r)+\Delta h_1\sum_{L_1L_2L_3}L_1^2=0,$$

$$\frac{\Delta h_1}{h_1}=-\frac{\sum L_1^2f(r)}{\sum L_1^2}\quad\text{or}\quad\frac{\Delta d}{d}=\frac{\sum L_1^2f(r)}{\sum L_1^2}.$$

It can be seen from this formula that the relative change in lattice constant depends primarily on the long-range effect of the distortion rather than upon the effects near the origin since each summation involves L_1^2. In other words, the result depends mostly on the form of $f(r)$ for large r values and is nearly independent of $f(r)$ at small r values (unless the crystallite contains only a few atoms, say less than 10^2). The theory of elasticity indicates that under quite general conditions (isotropic media in which Cauchy's relations hold) the strain $\Delta r/r$ varies as r^{-3}. This suggests that the appropriate form of $f(r)$ is $A(a/r)^3$ where A is a constant and $a=|\mathbf{a}_j|$. The subsequent discussion is based on the assumption that $f(r)$ is of this form; for a more exact treatment an additive constant must be included in $f(r)$, as shown in Appendix 5.

The results of summations and integrations for cubical crystallites containing $N=5^3$ to $N=30^3$ atoms indicate that

$$\Delta d/d=2.35\ (4\pi A/3N),$$

whereas for spherical crystallites the coefficient is found to be slightly larger:

$$\Delta d/d=2.5\ (4\pi A/3N).$$

The reason for writing the result in this form is that the macroscopic linear expansion is exactly $\Delta l/l=4\pi A/3N$, as shown in Appendix 2, and comparison of the two expansions may be readily made. For this idealized case the results are thus seen to differ by more than a factor of two.

That this result is qualitatively correct may be seen from the fact that the x-rays essentially average the interplanar spacings throughout the crystallite as the formula for $\Delta d/d$ shows, but the macroscopic linear expansion depends only on the displacement of the outermost atoms of the crystallite which are smaller than average since $f(r)$ is assumed to decrease with r.

II. ASYMMETRIC CRYSTALS AND EFFECT OF RANDOMLY DISTRIBUTED DEFECTS

In any attempt to observe the effects of such strains experimentally one must use specimens of macroscopic proportions having a large number of defects in un-

known positions. If a completely random distribution of identical defects could be obtained, say by heating the crystal and then quenching it to prevent coagulation of the defect centers, an estimate of the expected effect could be obtained directly from the results for spherical crystallites by applying appropriate weighting factors computed from the probability distribution.[1] Thus, let a spherical surface be constructed about each defect in the crystal, the diameter of the sphere being equal to the distance to the nearest defect. The probability that the closest defect is in the range dr at a given distance r is given by

$$\omega(r)dr=4\pi r^2 n[\exp(-4\pi r^3 n/3)]dr,$$

where n is the average number of defects per unit volume. From this function the relative number of spheres of each diameter can be computed, the contribution to the scattering can be weighted in accordance with the volume of the sphere, and an average $\Delta d/d$ can be found. Admittedly, a large fraction of the volume of the crystal is omitted completely by this procedure, but even if the contribution of the portions outside the spheres is assumed to be zero, the average $\Delta d/d$ is only about 30 percent lower than the value for a single sphere containing as many atoms as the average number of atoms per defect. It therefore appears that no large correction is needed, and the value of $\Delta d/d$ will still be much greater than $\Delta l/l$, which is completely independent of the distribution under these assumptions.

A second way of estimating the effect of the random distribution of defects upon $\Delta d/d$ is to visualize the randomness as a perturbation of the completely uniform distribution. Thus in a rectangular crystallite containing n defects per unit volume, suppose that the defects are arranged at equal intervals of $n^{-\frac{1}{3}}$ and that cubes having edges $n^{-\frac{1}{3}}$ are constructed about each defect. A certain amount of randomness can now be introduced by allowing each defect to move about in its cubical cell; the average distance between defects is only $0.55n^{-\frac{1}{3}}$, so that an average displacement of each defect from the center of its cube might be taken as $\frac{1}{2}n^{-\frac{1}{3}}$ for a crude approximation. Restricting the displacements to the three basic lattice vector directions $\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3$, the $\Delta d/d$ factor for each cell can be computed and a new average found. The calculations are more complicated as the symmetry properties required in the previous formula are not satisfied; approximations were used in the cases of displacements of the defects along the $\mathbf{a}_2$ and $\mathbf{a}_3$ directions but these are not believed to affect the result a great deal. The new average is found to be about 17 percent lower than for a single symmetrical cube of volume n^{-1}. From this and the previous result it is concluded that a reduction of about 20 percent in $\Delta d/d$ is to be expected.

Neither method of estimating the effect of the random distribution of the defects takes account of the fact that atoms near two (or more) defects undergo simultaneous displacements which may be comparable in magnitude. In order to calculate the change in the positions of the diffraction maxima in this general case, more elaborate formulas must be used. To determine the order of magnitude of this combination of displacements upon $\Delta d/d$, the case of a rectangular crystallite having the form of two adjoining cubes, each containing one defect at its center, was computed assuming that the line joining the two defects was parallel to one of the basic lattice vectors. The average value of $\Delta d/d$ was found (from the formulas in Appendix 1) to be only about 10 percent lower than that for a single cubical crystallite having one-half as many atoms and with the defect at the center of the cube. Thus, the neglect of the displacements due to all but one of the defects is not thought to invalidate the previous conclusion that $\Delta d/d$ is nearly twice as great as $\Delta l/l$. Allowing a 20 percent reduction, then

$$\Delta d/d=2.0(4\pi A/3N),$$

where N is the number of atoms per defect.

III. APPLICATION TO ACTUAL CRYSTALS

(A) AgBr

The results of Berry[2] and Strelkow[3] on AgBr indicate that $\Delta l/l$ and $\Delta d/d$ are virtually the same above 350°C. From the previous calculation the conclusion follows that Schottky defects certainly must occur in this temperature range since, if only Frenkel defects occur, there would be a factor of two between the two expansions. In the case of Schottky defects there is an additional contribution to the macroscopic expansion due to the diffusion of the displaced ions to the surface; and, hence, as shown in Appendix 3,

$$\Delta l/l=(1+4\pi A)/3N,$$

where N is the number of atoms per Schottky defect. However, the effect of the new surface layers upon $\Delta d/d$ is quite negligible. The experimental results indicate that $\Delta d/d$ and $\Delta l/l$ are to be equated so that if only Schottky defects occur:

$$2.0\ (4\pi A/3N)=(1+4\pi A)/3N,$$

which is satisfied if $A\sim 1/4\pi \doteq 0.08$.

From the form of the displacement function

$$\Delta\mathbf{r}=\mathbf{r}f(r)=\mathbf{r}A(a/r)^3,$$

it can be seen that the constant A denotes the fraction of the lattice spacing through which nearest neighbors adjacent to the defect are displaced in the ideal case. This is only an estimate in the case of an ionic crystal, such as AgBr, since ions of opposite sign are involved

[1] S. Chandrasekhar, Rev. Modern Phys. **15**, 1 (1943).

[2] C. R. Berry, Phys. Rev. **82**, 422 (1951).

[3] P. G. Strelkow, Phys. Z. Sowjetunion, **12**, 73 (1937).

and $f(r)$ is known to be irregular near a vacancy. From the calculations of Mott and Littleton[4] the relative displacement of nearest neighbors at a cation or anion vacancy in NaCl is thought to be about 0.07, so that the average A computed for AgBr by the above procedure is correct at least in order of magnitude. Of course, if a fraction α of the defects are Frenkel defects, then, as shown in Appendix 3, the average A value would have to be smaller in the ratio $(1-\alpha)/(1+\alpha)$ in order to explain the observed results. Thus, if $\frac{1}{3}$ of the defects are Frenkel type, the average value of A would be only 0.04 and the value decreases rapidly as α increases. The approximate agreement of A to the calculated values for $0<\alpha<\frac{1}{3}$ thus seems to indicate that Schottky defects predominate in AgBr above 350°C.

(B) KCl

The observations of Burstein, Oberly, and Davisson[5] of the effect of pressure upon the color center absorption in various alkali halides led them to conclude that there is a displacement of the ions near an F center toward the vacant site. This suggests that there may be an observable change in x-ray lattice constant when the concentration of F centers is approximately 10^{18} cm^{-3}. The concentration of scattering centers in the perfect KCl lattice is of the order of $10^{24}/30$, so that the number of atoms per defect is about

$$N=10^{24}/(30\times10^{18})=\tfrac{1}{3}\times10^{5}\text{ cm}^{-3}.$$

Hence, the expected increase in lattice constant is

$$\Delta d/d=2(4\pi A/3N)=\tfrac{2}{3}(-1)(3\times10^{-5})=-2\times10^{-5},$$

assuming that $A=-\frac{1}{4}\pi$. F center concentrations of 5×10^{18} cm^{-3} would then produce a relative change in lattice constant of approximately one part in ten thousand, which is certainly detectable and with proper precautions might be measured with 10 percent accuracy. The interpretation of the results of such an experiment would be less ambiguous than those for AgBr as the F centers can be almost completely dispersed by proper heat treatment (quenching the crystal by the method of Scott and Bupp[6]), and by additive coloring a predominance of F centers is assured so that the effect of other types of defects may be neglected safely. If sufficient accuracy is attained, there is the possibility that the results of the x-ray measurements could be compared to density measurements[7] to determine the displacement constant since the two quantities depend upon it in different ways (see Appendix 4):

$$\Delta d/d\doteq2(4/3\pi A)/N,\quad \Delta\rho/\rho\doteq-(4\pi A+1)/N.$$

Both vary as N^{-1} so that the ratio $\Delta d/d\div\Delta\rho/\rho$ depends upon A alone.

The observations of Pick and Weber[8] of density changes in KCl due to small concentrations of $CaCl_2$ and $SrCl_2$ indicate that lattice vacancies occur, although not quite in accordance with the substitution theory in which each Ca and Sr impurity ion produces a K ion vacancy. The discrepancy may possibly be explained by an over-all decrease in the lattice constant due to the defects. The Goldschmidt radii of both Ca and Sr ions are definitely less than for K ions, and the resultant shrinking about these ions may offset the expansion occurring at the vacancies. Further experiments with $BaCl_2$ impurities might shed some light on this question, as the radius for the Ba ion is only about 2 percent greater than that for the K ion. Thus, with $BaCl_2$ as an impurity, the principal effect upon the lattice constant could be ascribed to the vacancies alone and the interpretation of the results greatly simplified.

(C) LiF

Measurements of the thermal expansion of several alkali halides have been reported by A. Gott[9] who found that the macroscopic linear expansion was larger than the x-ray lattice constant expansion in every case. In LiF at 350°C, the difference in the two relative expansions is of the order of 10^{-3}. Assuming that only Schottky defects occur (see Appendix 3),

$$\frac{\Delta l}{l}-\frac{\Delta d}{d}=\frac{1}{3N}(1-4\pi A)\sim10^{-3}.$$

From this it can be seen that $N<10^{3}$ if $|A|\sim\frac{1}{4}\pi$. From diffusion and conduction experiments,[10] one expects $N\sim10^{5}$ (in rock salt) at this temperature, so that it seems unlikely that the difference in the two expansions can be explained by Schottky defects alone. More measurements of these crystals at high temperatures might clarify the interpretation of these results.

(D) Mixed Crystals

The theory does not seem to be applicable to the case of mixtures of ionic crystals. C. R. Berry[11] has pointed out that the data on the system KCl-KBr indicate a ratio $\Delta d/d\div\Delta l/l\sim1$ instead of 2.0. However, for large impurity concentrations ($\sim$10 percent) upon which this ratio is based, the theory may not apply; moreover, the calculation of $\Delta l/l$ is made from density changes and assumes that the mixed crystals are perfect—i.e., free of voids which would make $\Delta l/l$ erroneously large but would not affect the x-ray measurements of $\Delta d/d$. For these reasons it is believed that the experimental observations do not

[4] N. F. Mott and R. W. Gurney, *Electronic Processes in Ionic Crystals* (Oxford University Press, New York), p. 61.
[5] Burstein, Oberly, and Davisson, Bull. Am. Phys. Soc. **26**, No. 6, Paper 03.
[6] A. B. Scott and L. P. Bupp, Phys. Rev. **79**, 341 (1950).
[7] Estermann, Leivo, and Stern, Phys. Rev. **75**, 627 (1949).
[8] H. Pick and H. Weber, Z. Physik **128**, 409 (1950).
[9] A. Gott, Ann. Phys. Lpz. **41**, 520 (1942).
[10] Mapother, Crooks, and Maurer, J. Chem. Phys. **18**, 1231 (1950).
[11] Private communication based on observations by: F. Oberlies, Structurbericht **2**, 214 (1928–32); G. Tammann and W. Krings, Z. anorg. u. allgem. Chem. **130–32**, 242 (1923–4).

prove that the theory is wrong; in any case it would be desirable to have more data on such systems.

APPENDIX 1. GENERAL FORMULA FOR RELATIVE CHANGE IN LATTICE CONSTANT AS MEASURED FROM X-RAY DIFFRACTION MAXIMA

The relative amplitude of the wave scattered from an array of identical atoms at positions $\mathbf{R}$ is given by the summation over the $\mathbf{R}$ vectors:

$$E=\sum_{R}\exp\left[2\pi i\left(\frac{\mathbf{S}-\mathbf{S}_0}{\lambda}\cdot\mathbf{R}\right)\right].$$

In the undistorted crystal the $\mathbf{R}$ vectors are given by

$$\mathbf{R}=\mathbf{r}=L_1\mathbf{a}_1+L_2\mathbf{a}_2+L_3\mathbf{a}_3,$$

and at the diffraction maxima the unit vectors $\mathbf{S}_0$ and $\mathbf{S}$ denoting the directions of the incident and scattered waves satisfy the condition

$$\frac{\mathbf{S}-\mathbf{S}_0}{\lambda}=h_1\mathbf{b}_1+h_2\mathbf{b}_2+h_3\mathbf{b}_3=\mathbf{B},$$

where the vector set $\mathbf{b}_j$ is the reciprocal of the set $\mathbf{a}_j$ and the sets L_j and h_j are integers. In the distorted crystal the positions of the atoms may be denoted by

$$\mathbf{R}=\mathbf{r}+\Delta\mathbf{r}=(L_1+\Delta L_1)\mathbf{a}_1+(L_2+\Delta L_2)\mathbf{a}_2+(L_3+\Delta L_3)\mathbf{a}_3,$$

where ΔL_j may be considered as functions of $\mathbf{r}$ alone and $\Delta L_j\ll 1$ for all r. To determine the values of $\mathbf{S}-\mathbf{S}_0/\lambda$ for which the intensity of the scattered x-rays from the distorted crystal is a maximum, the deviation from the reciprocal lattice vector of the perfect lattice may be assumed to be small and expressible in the form $\mathbf{\Delta B}$, so that

$$\frac{\mathbf{S}-\mathbf{S}_0}{\lambda}=\mathbf{B}+\mathbf{\Delta B}=(h_1+\Delta h_1)\mathbf{b}_1+(h_2+\Delta h_2)\mathbf{b}_2+(h_3+\Delta h_3)\mathbf{b}_3.$$

The expression for the relative amplitude of the scattered wave then becomes

$$E=\sum_{r}\exp\{2\pi i[(\mathbf{B}+\mathbf{\Delta B})\cdot(\mathbf{r}+\mathbf{\Delta r})]\}$$

$$\doteq\sum_{r}\exp\{2\pi i[\mathbf{\Delta B}\cdot\mathbf{r}+\mathbf{B}\cdot\mathbf{\Delta r}]\}=\sum_{r}\exp\{2\pi i\delta(r)\},$$

when the factor $\exp 2\pi i(\mathbf{B}\cdot\mathbf{r})$ is omitted since it is unity and the second-order term $\mathbf{\Delta B}\cdot\mathbf{\Delta r}$ is neglected and

$$\delta(r)=\Delta h_1L_1+\Delta h_2L_2+\Delta h_3L_3+h_1\Delta L_1+h_2\Delta L_2+h_3\Delta L_3.$$

Expanding the exponential in its power series:

$$E=\sum_{r}\left(1+2\pi i\delta-\frac{1}{2!}[2\pi\delta]^2\right).$$

Only the first three terms are used assuming that $\delta(r)$ is small over the entire range of the summation. The relative intensity of the scattered wave is then of the form

$$EE^*=\left\{\sum_{r}\left(1-\frac{1}{2!}[2\pi\delta]^2\right)\right\}^2+\{\sum_{r}2\pi\delta\}^2,$$

and the values of Δh_j for which the intensity is maximum satisfy the three conditions simultaneously:

$$\frac{\partial(EE^*)}{\partial(\Delta h_j)}=0,\quad j=1,2,3.$$

Considering the first derivative as an example,

$$\frac{\partial(EE^*)}{\partial(\Delta h_1)}=2\left\{\sum_{r}\left(1-\frac{1}{2!}[2\pi\delta]^2\right)\right\}\times\{\sum_{r}(0-[2\pi\delta]2\pi L_1)\}+2\{\sum_{r}2\pi\delta\}\{\sum_{r}2\pi L_1\}=0.$$

Assuming $2\pi^2\delta^2\ll 1$,

$$-N\sum_{r}L_1\delta+(\sum_{r}L_1)(\sum_{r}\delta)=0,$$

where N denotes the total number of atoms in the crystal given by

$$N=\sum_{r}(1).$$

Regarding Δh_1, Δh_2, Δh_3 as unknowns independent of the summations the equation may be rearranged for computation as shown below; the second and third equations are obtainable from the second and third conditions:

$$\begin{aligned}&\Delta h_1\{(\textstyle\sum L_1)^2-N\sum L_1^2\}\\&\quad+\Delta h_2\{(\textstyle\sum L_1)(\sum L_2)-N\sum(L_1L_2)\}\\&\quad+\Delta h_3\{(\textstyle\sum L_1)(\sum L_3)-N\sum(L_1L_3)\}\\&=N\textstyle\sum[L_1(h_1\Delta L_1+h_2\Delta L_2+h_3\Delta L_3)]\\&\qquad-(\textstyle\sum L_1)(\sum[h_1\Delta L_1+h_2\Delta L_2+h_3\Delta L_3]).\end{aligned}\quad(1)$$

$$\begin{aligned}&\Delta h_1\{(\textstyle\sum L_1)(\sum L_2)-N\sum L_1L_2\}\\&\quad+\Delta h_2\{(\textstyle\sum L_2)^2-N\sum(L_2^2)\}\\&\quad+\Delta h_3\{(\textstyle\sum L_2)(\sum L_3)-N\sum(L_2L_3)\}\\&=N\textstyle\sum[L_2(h_1\Delta L_1+h_2\Delta L_2+h_3\Delta L_3)]\\&\qquad-(\textstyle\sum L_2)(\sum[h_1\Delta L_1+h_2\Delta L_2+h_3\Delta L_3]).\end{aligned}\quad(2)$$

$$\begin{aligned}&\Delta h_1\{(\textstyle\sum L_1)(\sum L_3)-N\sum(L_1L_3)\}\\&\quad+\Delta h_2\{(\textstyle\sum L_2)(\sum L_3)-N\sum(L_2L_3)\}\\&\quad+\Delta h_3\{(\textstyle\sum L_3)^2-N\sum L_3^2\}\\&=N\textstyle\sum[L_3(h_1\Delta L_1+h_2\Delta L_2+h_3\Delta L_3)\\&\qquad-(\textstyle\sum L_3)(\sum[h_1\Delta L_1+h_2\Delta L_2+h_3\Delta L_3]).\end{aligned}\quad(3)$$

In the case of an (h_100) reflection from a cubical crystallite having a symmetric distortion relative to the center so that

$$\sum L_j=\sum\Delta L_j=0,\quad j=1,2,3$$

and

$$\sum L_j L_k = 0, \quad j \neq k$$

(when the origin is taken at the center), the first equation reduces to

$$\Delta h_1\{-N\sum L_1^2\} = N h_1 \sum L_1 \Delta L_1,$$

or

$$\frac{\Delta d}{d} = -\frac{\Delta h_1}{h_1} = \frac{\sum L_1 \Delta L_1}{\sum L_1^2}.$$

The second and third equations yield $\Delta h_2 = \Delta h_3 = 0$ since sums of the form $\sum L_j \Delta L_k = 0$ when $j \neq k$. These results justify the assumptions made in the text for this case. To make the results completely identical one has only to remember that in the special case of radial distortion it was assumed that

$$\mathbf{\Delta r} = \mathbf{r} f(r) = f(r)(L_1\mathbf{a}_1 + L_2\mathbf{a}_2 + L_3\mathbf{a}_3),$$

whereas in the general case

$$\mathbf{\Delta r} = \Delta L_1 \mathbf{a}_1 + \Delta L_2 \mathbf{a}_2 + \Delta L_3 \mathbf{a}_3.$$

In the special case that $\Delta L_1 = L_1 f(r)$, then

$$\frac{\Delta d}{d} = \frac{\sum L_1^2 f(r)}{\sum L_1^2}.$$

A fairly simple result is also obtainable for the case of $(h_1 00)$ reflections from a crystal in which the origin of the distortion is displaced along the $\mathbf{a}_1$ direction so that $\sum L_1 \neq 0$ and $\sum \Delta L_1 \neq 0$. However,

$$\left.\begin{aligned} \sum L_1 L_j &= 0 \\ \sum L_j \Delta L_1 &= 0 \end{aligned}\right\} \quad j = 2, 3,$$

so that $\Delta h_2 = \Delta h_3 = 0$ as before and

$$\frac{\Delta d}{d} = -\frac{\Delta h_1}{h_1} = \frac{(\sum L_1)(\sum \Delta L_1) - N\sum(L_1 \Delta L_1)}{(\sum L_1)^2 - N\sum L_1^2},$$

or

$$\frac{\Delta d}{d} = \frac{(\sum L_1)(\sum L_1 f(r)) - N\sum L_1^2 f(r)}{(\sum L_1)^2 - N\sum L_1^2}.$$

In the general case the three linear equations in Δh_j must be solved simultaneously. Keeping the equations in their most general form makes it possible to compute the effect of a large number of randomly distributed defects in a single crystal. Thus, the locations of the defects in the crystal could be determined by choosing sets of three numbers at random, the effect of each defect upon the position of each atom could be computed, and the sum of the displacements determined for each atom; from these sums the ΔL_j values could be obtained and all of the coefficients in the three equations (1–3) could be computed. Such a method of finding Δh_j and thence $\mathbf{\Delta B}$ would be too laborious for hand calculation, but machine methods are obviously possible and are being considered.

APPENDIX 2. EXPANSION OF A CRYSTAL CAUSED BY RADIAL DISTORTION

In the case of a radial distortion such that

$$\mathbf{\Delta r} = \mathbf{r} f(r) = \mathbf{r} A \left(\frac{a}{r}\right)^3,$$

the change in volume of a spherical crystallite of radius r_0 containing N atoms is

$$\Delta V = (4\pi r_0^2)[\Delta r(r_0)] = 4\pi r_0^2 [A a^3 r_0^{-2}] = 4\pi a^3 A,$$

so that

$$\Delta V/V = \frac{4\pi a^3 A}{N a^3} = \frac{4\pi A}{N}$$

and hence the relative linear expansion is approximately

$$\frac{\Delta l}{l} = \frac{1}{3}\frac{\Delta V}{V} = \frac{4\pi A}{3N}.$$

As the increment in volume is independent of the dimensions of the crystal, the effect of having more than one lattice defect in the same crystallite is simply to increase the expansion linearly with the defect concentration. Hence the same general equation holds for $\Delta l/l$, but N becomes the average number of atoms per defect.

If Frenkel defects occur there will in general be a different A coefficient for each constituent. In the case of ionic crystals, both the lattice vacancy and interstitial ion produce a swelling of the lattice so that the effects do not cancel; in any case the results can be described by an average A constant. If Schottky defects occur there is a quite different contribution to the expansion due to the ions which diffuse to the surface; the ions at the surface always increase the linear dimensions and if $|A| < \frac{1}{4}\pi$ the contribution of the surface layers will predominate over that due to the vacancies.

APPENDIX 3. COMPARISON OF $\Delta d/d$ AND $\Delta l/l$ FOR FRENKEL AND SCHOTTKY DEFECTS IN IONIC CRYSTALS

Owing to the fact that $\Delta d/d \sim 2\Delta l/l$ for any lattice defect, whether a vacancy or interstitial ion, any crystal having only Frenkel defects should have $\Delta d/d \sim 2\Delta l/l$. Assuming for simplicity that the A coefficients are the same for vacancies and interstitial ions and that the fraction of Frenkel defects is α, then there are αn Frenkel vacancies and $(1-\alpha)n$ Schottky vacancies per N' lattice sites where n is the total number of defects and $N = N'/n$ is the average number of lattice sites per defect. The number of interstitial ions is αn, and hence

$$\frac{\Delta d}{d} = 2\left(\frac{4\pi A}{3N'}\right)[2\alpha n + (1-\alpha)n] = 2\left(\frac{4\pi A(1+\alpha)}{3N}\right).$$

But

$$\Delta V=4\pi A a^3[2\alpha n+(1-\alpha)n]+(1-\alpha)na^3,$$

since $(1-\alpha)n$ ions diffuse to the surface from the Schottky defects. Hence

$$\frac{\Delta V}{V}=\frac{na^3}{N'a^3}[4\pi A(1+\alpha)+(1-\alpha)]$$

and

$$\frac{\Delta l}{l}=\frac{1}{3N}[4\pi A(1+\alpha)+(1-\alpha)].$$

In the special case $\Delta d/d=\Delta l/l$ then

$$4\pi A(1+\alpha)=(1-\alpha)$$

or

$$A=\frac{1}{4\pi}\frac{1-\alpha}{1+\alpha}.$$

APPENDIX 4. CHANGE IN DENSITY CAUSED BY F CENTERS PRODUCED BY ADDITIVE COLORING

In additively colored crystals there is a predominance of negative ion vacancies (F centers); for each vacancy formed there are two ions at the surface, and if the masses are m_1 and m_2 for the positive and negative ions, respectively, the increase in mass per F center is m_1 but the increase in volume is $4\pi Aa^3+2a^3$ due to the effect of the vacancy. The relative change in density for a crystal containing N' atoms and n F centers is then

$$\frac{\Delta\rho}{\rho}=\frac{\Delta(M/V)}{M/V}=\frac{V\Delta M-M\Delta V}{V^2}\frac{V}{M}=\frac{\Delta M-M(\Delta V/V)}{M}$$

$$=\frac{nm_1-N'(m_1+m_2/2)[(4\pi A+2)a^3n/N'a^3]}{[(m_1+m_2)/2]N'}$$

$$=\frac{m_1-[m_1+m_2/2](4\pi A+2)}{(m_1+m_2/2)N},$$

where as before $N=N'/n$.

When m_1 and m_2 are equal, then

$$\Delta\rho/\rho=-(4\pi A+1)/N.$$

Of course if $m_1=m_2$ and $A=-\frac{1}{4}\pi$ then $\Delta\rho/\rho$ vanishes; the change in x-ray lattice constant would still be observable, however.

APPENDIX 5. DISTORTION IN ISOTROPIC MEDIA

The exact solution for the radial displacement in a hollow sphere of elastic material obeying Cauchy's relations and having Young's modulus E is

$$\Delta r=\frac{1}{2E}\frac{p_1r_1^3-p_0r_0^3}{r_0^3-r_1^3}r+\frac{5}{8E}\frac{r_0^3r_1^3(p_1-p_0)}{r_0^3-r_1^3}\frac{1}{r^2},$$

where the internal surface has radius r_1 and is subject to pressure p_1; the outer surface is concentric with the inner one and has radius r_0 and pressure p_0.[12]

Assuming that the effect of a vacancy or interstitial atom is to create an internal pressure, that the external pressure is zero, and $r_0\gg r_1$,

$$\Delta r=\frac{1}{2E}p_1\frac{r_1^3r}{r_0^3}+\frac{5}{8E}p_1\frac{r_1^3}{r^2}=r\left[A\left(\frac{a}{r}\right)^3+B\right]=rf(r),$$

where $A=5p_1r_1^3/8Ea^3$ and $B=p_1r_1^3/2Er_0^3$. Thus the form for $f(r)$ assumed in the text is oversimplified, but correct in the power of r; the added constant introduces no difficulty in the calculation; the results for single spheres become

$$\frac{\Delta d}{d}=2.5A\frac{4\pi}{3N}+B \quad \text{and} \quad \frac{\Delta l}{l}=\frac{4\pi A}{3N}+B.$$

The ratio of B to A for this case is

$$\frac{B}{A}=\frac{4}{5}\left(\frac{a}{r_0}\right)^3=\frac{16\pi}{15N}, \quad \text{since} \quad \frac{4}{3}\pi r_0^3=Na^3,$$

so that $\Delta d/d\div\Delta l/l=33/18$ instead of $5/2$ for single spherical crystallites.

In crystals containing large numbers of defects distributed at random the method of calculation employed in the text indicates that a slight reduction in the first term for $\Delta d/d$ will persist; however, a relatively large correction in the value of the constant terms for the spheres constructed about each defect is to be expected since the external pressure at these surfaces is by no means zero, so that

$$B=\frac{1}{2Er_0^3}(p_1r_1^3-p_0r_0^3),$$

and B may nearly vanish. On account of this effect it is believed that the ratio 33/18 is too small and that the value $\Delta d/d\div\Delta l/l\sim 2.0$ derived in the text is most nearly correct

[12] A. E. H. Love, *Mathematical Theory of Elasticity* (Cambridge University Press, Cambridge, 1927), p. 142.

Bonding of Thin Films*

H. P. MEISSNER AND J. BYRNE
Chemical Engineering Department, Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received June 30, 1952)

Unsupported films of nitrocellulose of 0.25 micron and less in thickness bonded immediately and spontaneously in air to solid surfaces which they were made to touch. Bond strength increased with diminishing film thickness. Similar behavior was shown by films of gold, regenerated cellulose, polymethyl methacrylate, and gelatin. This bonding power of nitrocellulose films, which was apparently not attributable to static electricity or to mechanical joint formation, was lost while immersed in water, benzene, or hexane, but was regained upon removal and drying.

THIS study was undertaken following the accidental observation that films of nitrocellulose less than about one-quarter of a micron in thickness showed a bonding behavior not encountered with thicker films or with nitrocellulose in bulk. That is, it was found that such thin nitrocellulose films stuck spontaneously and immediately to any solid surface with which they were brought into contact. The resultant bond was sometimes so strong that the film could not be peeled away without tearing. Substantially thicker films of nitrocellulose showed no such bonding behavior even when pressed upon a surface with 1000 lb force per square inch. It is the object here to describe an experimental investigation of this phenomenon.

PROCEDURE

Thin films of various thicknesses for each of the materials were made and tested to determine their bonding properties. The materials investigated, namely, nitrocellulose, polymethyl methacrylate, cellulose, gelatin, and gold, are characterized in Table I. After formation, the films were stored in a covered box to protect them from dust.

Film Formation

A. Nitrocellulose

A few drops of a 3 percent by weight solution of nitrocellulose in technical grade amyl acetate were placed on the surface of distilled water in a large beaker. A film formed upon the water surface as the amyl acetate evaporated and dissolved in the distilled water. A cylinder made of aluminum pipe, roughly $\frac{1}{4}$-in. diameter and 1-in. long, was next lifted up through this film so that the film closed one end of this cylinder. This mounted film was dried for 24 hours in a ventilated oven at 95°C.

B. Polymethyl Methacrylate

A 5 percent solution by weight of polymethyl methacrylate was prepared, using a solvent containing approximately 10 percent amyl acetate and 90 percent benzene, the proportions being adjusted to give a clear, smooth film. Films were cast, mounted, and dried as in the case of nitrocellulose.

C. Cellulose

Cotton linters were dissolved in 60 percent aqueous benzyltrimethylammonium hydroxide, forming a 5-weight percent solution. A few drops of this solution were placed on the surface of a 10 percent water, 90 percent glycerin mixture, whereupon a film formed. Several hours later, these films were strong enough to be mounted over a $\frac{1}{4}$-inch hole in an aluminum plate, gently washed for 12 hours in distilled water, and dried in an air-conditioned room at 23°C and 65 percent relative humidity.

D. Gelatin

Gelatin was dissolved in hot water to form a 3 percent solution. This solution was cooled to within a few degrees of the gelation point, whereupon a loop of thin wire was drawn up through the solution and the resulting film was allowed to dry at room temperature. This film was then floated off the wire loop onto the surface of distilled water maintained at room temperature (room temperature being well below gelation temperature). The film was mounted on an aluminum plate, and then dried in an air-conditioned room at 23°C, and the desired humidity.

E. Gold

Commercial gold leaf, about 2.5 microns in thickness, was mounted by being clamped between glass rings each roughly 1 inch in diameter. By dissolving the excess gold in aqua regia, it was possible to obtain gold films of thicknesses from 0.15 to 1.3 microns.

Measurement of Thickness

The thickness of all but the gold films was measured with a Jamin interferometer[1] using the following formula:

$$(n-1)(z)=(\lambda)(N),$$

* Based in part on a thesis presented by J. Byrne in partial fulfillment of the requirements for the degree of Doctor of Science at the Massachusetts Institute of Technology.

[1] W. E. Williams, *Applications of Interferometry* (Metheun and Company, Ltd., London, 1929), p. 42.

where n is the index of refraction, z is thickness in millimicrons, λ is the wavelength of light used in millimicrons, and N is the number of interference fringes determined by the Jamin interferometer. It was assumed that the index of refraction for these materials in bulk was the same as in thin films and the following values were used: nitrocellulose[2] 1.51, gelatin[3] 1.525, cellulose[4] 1.54, and polymethyl methacrylate[5] 1.41. The reproducibility of the thickness measurement on any one material, as calculated from the above relation and values, was found to be plus or minus 0.02 micron.

The gold film thickness was measured with a Sheffield Visual Gage, which utilizes a combination of optical and mechanical amplification to provide readings which have a precision of plus or minus 0.05 micron.

Bond Testing

A qualitative test of bonding behavior involved bringing the mounted film into contact with a solid surface or a second film. Development of a bond was judged by whether or not a detectable force was required to peel the film manually from the solid surface. Variation in the tension of the mounted films made it impractical to place this test upon a quantitative basis. Moreover, quantitative results of any peeling test are particularly difficult to translate into useful terms, since the results are complicated by film stiffness, thickness, and the like.

A shear test was finally adopted for the bond evaluations reported here. This test involved bringing the flat end of a tiny metal cylinder, herein called the "probe," very gently into contact with the film to be tested, as shown in Fig. 1. This test could be carried out with the film and probe in air, or totally immersed in a liquid. The flat end of the probe was kept polished with Crocus cloth. Due care was exercised to maintain this flat end parallel to the plane of the film surface at all times, and to assure that no significant force was applied normal to the plane of the bond. A force was next imposed upon the probe tending to move it in a direction parallel to the film surface. The minimum force required to move the probe, in other words to break the bond between the probe and the film in shear, was then determined. The probe was attached to an extension of one arm of an analytical balance, and a shearing force could be applied and increased at will by adding weights to the pan connected to the other arm of the balance. The reproducibility of measurement of this breaking load was found to be plus or minus 0.02 gram when testing a given area on one film.

[2] Emil Ott, editor, *Cellulose and Cellulose Derivatives* (Interscience Publishers, Inc., New York, 1943), p. 1070.
[3] C. D. Hodgman, editor, *Handbook of Chemistry and Physics* (Chemical Rubber Publishing Company, Cleveland, 1948), thirtieth edition, p. 2224.
[4] P. H. Hermans, *Contributions to the Physics of Fibers* (Elsevier Publishing Company, Inc., Houston, 1946), p. 118.
[5] Reference 3, p. 910.

Critical Bonding Thickness

The maximum thickness at which bonding could be detected by this test procedure is herein called the "critical bonding thickness." This determination involved testing a series of films, each thicker than the preceding by an increment of roughly 0.03 micron, and ranging from a thickness of 0.01 micron for the thinnest film to 0.6 micron for the thickest. In this test an arbitrary shearing force of 0.02 gram was imposed on the probe, as being the smallest force whose action could be conveniently observed. It was assumed that a bond had formed when the probe remained stationary. An equivalent determination of critical thickness is the extrapolation of a shear-test curve of the type shown in Fig. 2.

RESULTS

The effect of film thickness upon the load required to break a bond in air between a brass probe and nitrocellulose films is shown in the typical curve, presented in Fig. 2. It is evident that no detectable bond formed down to a film thickness of about 0.25 micron, at which point a bond appeared to develop, and below which the

TABLE I. Critical thickness of bonding.

Probe tests	Critical thickness of bonding microns	Temperature °C
Nitrocellulose		
Collodion, mallinkrodt "parlodion"	0.22–0.24	23 to 85
Hercules, 11.97 percent N, viscosity 17 seconds	0.22–0.24	23 and 70
Hercules, 12.05 percent N, viscosity 850 seconds	0.22–0.24	23 and 70
Cellulose		
Regenerated from a solution of cotton linters in benzyl-trimethyl ammonium hydroxide, 65 percent relative humidity	0.17–0.25	23
Polymethyl methacrylate		
Rohm and Haas; molecular weight of around 10^6 by viscosity measurement	0.19–0.23	23
Gelatin		
Atlantic Brand "super-clarified" (General Foods Corporation)		
relative humidity 10 to 65 percent	0.14–0.16	23
relative humidity 70 percent	0.16–0.18	23
relative humidity 80 percent	0.21–0.25	23
relative humidity 90 percent	0.27–0.33	23
Gold		
23 karat, purchased in sheets 0.0001 inch and then reduced to final thickness by acid leach	0.2 –0.25	23
Lamination tests		
Laminated nitrocellulose films "parlodion" total thickness of laminate	0.23–0.25	23
Two film tests "parlodion" thickness of thinner film	0.2 –0.25	23

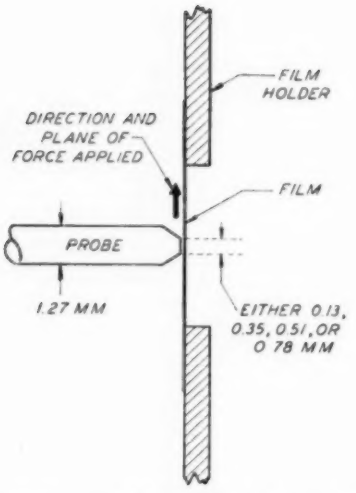


FIG. 1. Schematic diagram of apparatus for testing bond strength.

breaking load increased rapidly with decreasing thickness. The critical bonding thickness in this case was therefore somewhere around 0.25 micron. No bond strength data are presented in Fig. 2 at thicknesses below 0.1 micron because such very thin films usually tear before the load can be raised enough to break the bond in shear. Temperature and average molecular weight appear to have no effect upon the curve obtained, within the range of these variables studied.

For any given film, the load required to break a bond was found to vary more nearly with the diameter of the brass probe than with its area. This is evident from Table II, where there are presented breaking load data as read from curves similar to those from Fig. 2 for brass probes of different diameters bonded to nitrocellulose films. Some uniformity of values is evident for stress per unit diameter (load in grams)/(probe diameter in mm), while no uniformity is apparent for values for the ratio (load in grams)/(probe area in sq mm).

The critical bonding thicknesses for various materials, determined in part by Bowman,[6] Cohen,[7] Duhnkrack,[8] and Kunstenaar[9] are presented in Table I. These results were found to be independent of whether the probe was made of brass, nitrocellulose, or steel. The range of values reported is from the thickest film which always gave a definite bond to the thinnest film which never bonded. This spread appears partly due to the difficulty of obtaining uniform films, and is relatively large with cellulose and gold because these films were least uniform. There is also some evidence that the conditions of drying the films contribute to the scatter of the data.

Also presented in Table I are findings for nitrocellulose laminates. The critical bonding thickness of laminates containing two to four component films was the same as that of monofilms. Once formed, a laminate did not have bonding capacity unless it was thinner than 0.25 micron. The component films recovered their original bonding power upon being peeled away from the laminate. Similarly, thin films lost their capacity for further bonding after bonding to the surface of any relatively thick solid. They also recovered their original bonding power upon being peeled free.

Thin films showed a complete loss of bonding power while immersed in nonsolvent liquids. For example, nitrocellulose films having a thickness well under 0.25 micron did not bond under water, *n*-heptane or benzene. Upon removal from water, the nitrocellulose films again bonded upon drying or upon being shaken to remove surface water. On the other hand, films dipped into hydrocarbon solvents recovered their bonding capacities only slowly after immersion. Drying for about one hour in air was required before restoration was complete.

No special attempts were made to protect films after formation other than storage in air in a covered box. Some diminution of bonding power occurred with aging, in that films one month old formed bonds which were only two-thirds as strong as when new.

DISCUSSION OF RESULTS

Bond Testing

The bond test used here is not entirely satisfactory, since it does not cause simple failure in shear. That is, part of the film under the probe face stretches and slips before break occurs. This may account for the fact that breaking load is approximately constant when expressed in grams per unit of diameter. It therefore appears that this test method cannot give more than a qualitative indication of bond strength. It is hoped that further work will develop test procedures whereby the true tensile or shearing strength of bonds formed by these thin films can be determined.

TABLE II. Bond strength tests.

Nitrocellulose, Hercules Powder Company, viscosity 850 seconds, brass probe, at 23°C and 50 percent relative humidity

Thickness micron	Probe diameter mm	Breaking[a] load g	Linear stress g force/mm diameter	Area stress g force/sq mm of probe area
0.20	0.13	0.03	0.23	2.3
0.20	0.35	0.08	0.23	0.82
0.20	0.51	0.15	0.3	0.75
0.20	0.78	0.16	0.2	0.33
0.15	0.13	0.15	1.15	11.5
0.15	0.35	0.35	1.0	3.6
0.15	0.51	0.55	1.1	2.8
0.15	0.78	0.58	0.75	1.2
0.10	0.13	0.55	4.2	42[b]
0.10	0.35	1.00	2.8	10
0.10	0.51	1.20	2.4	6.0
0.10	0.56	1.35	2.4	2.8

[a] These data read from curves like those of Fig. 2.
[b] Highest value of shearing stress found was 90 g force/sq mm.

[6] R. J. Bowman, Master of Science thesis, M.I.T., 1951.
[7] E. S. Cohen, Master of Science thesis, M.I.T., 1951.
[8] G. B. Duhnkrack, Bachelor of Science thesis, M.I.T., 1951.
[9] M. Kunstenaar, Master of Science thesis, M.I.T., 1950.

Despite the limitations of the test methods used, it appears that a bond did in fact form between the probe and the thin films tested. These bonds differed from those formed by materials in bulk with respect to the effect of temperature, pressure, and the rate of bond formation. That is, Meissner and Merrill[10] found that materials brought into contact with a smooth surface would adhere only at temperatures above their second order transition temperatures. Thus, bonding of bulk nitrocellulose occurs only above its second-order transition temperature of 55°C. On the other hand, thin films of nitrocellulose bonded equally well at 23°C, and 70°C, and 85°C as shown in Fig. 2 and Table II. Again, even at temperatures considerably above 55°C, bulk nitrocellulose will form bonds only after the application of pressure for a finite time. Roughly speaking, attainment of a given bond strength at a given temperature in such a bulk system requires that the product of pressure and application time must equal a certain value, smaller values of this product resulting in lower bond strength. Again by way of contrast, thin films bond at all temperatures upon the lightest contact and, as far as discernible, bond strength is immediately attained. It should be noted, however, that the effect of pressure upon film bonding has not yet been explored.

The development of a mechanical, or interlocking, type of bond between the film and the probe might be a possible explanation of the thin film bonding described here. Two types of such bonds are imaginable: first, a bond caused by the penetration of the probe into the film with concomitant folding of the film around the end of the probe; second, a bond formed by minute spines on the probe surface which penetrate into the film. However, microscopic examination failed to reveal evidence either of film folding, as just described, or of spines on the probe. Moreover, these explanations fail to account for the bonding detected between films in lamination experiments, where no such mechanical interlocking could have developed. These qualitative lamination experiments resulted in the same value for the critical thickness as experiments using a probe. Finally, the failure of a bond to form between probe and film while totally immersed in a liquid indicates that the joint formed is not mechanical. It therefore appears that bonds formed by these thin films cannot be attributed to mechanical causes.

Early in the work it appeared that electrostatic charges might have an important effect upon the results obtained. The following findings were taken as evidence that electrostatic charges were not of primary importance: First, the results were reproducible while electrostatic phenomena are notoriously erratic; second,

[10] H. P. Meissner and E. W. Merrill, Modern Plastics, p. 104 (April, 1949).

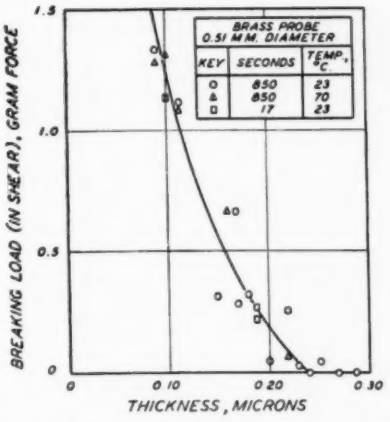


FIG. 2. Relationship between breaking load and film thickness for nitrocellulose films tested at 23°C and 70°C with a brass probe 0.51 mm in diameter.

the films were retested after exposure for 24 hours to an ionized atmosphere with no change in the findings; third, the films were deliberately charged with no change in the findings; and fourth, the gold-film experiments were made with both probe and film electrically grounded.

It has been proposed that adsorbed gases and liquids present on the surfaces of these films were responsible for the bonding behavior observed. Since the amount of material adsorbed on a surface is presumably independent of film thickness, an explanation of bonding based primarily on surface impurities is inconsistent with the experimental finding that bond strength varies inversely with film thickness.

These films, being extremely thin, are easily deformed and so can nestle closely into the uneven surface of any solid which they touch. However, there appear to be two reasons why film bonding behavior cannot be attributed to the capacity to deform which results from thinness. First, it appears unlikely that the capacity to deform would be seriously affected by small quantities of hydrocarbons, yet these thin films temporarily lost bonding power after being immersed in hydrocarbons. Second, the capacity to deform is determined by properties such as the moduli of the film material. The five materials used here have very different physical properties, yet they showed an equal critical bonding thickness. This suggests that their bonding behavior is not primarily attributable to their capacity to deform.

A qualitative survey, not yet complete, suggests that the thin films described here will bond to most solids, regardless of composition. Thin films of many other materials may be found to behave similarly, but the temporary failure of nitrocellulose films to bond after immersion in hydrocarbons shows that not all thin films will bond. The general limitations and explanations for thin film bonding can be determined only after further investigation.

Theory of Electrically Short Transmitting and Receiving Antennas*

RONOLD KING
Cruft Laboratory, Harvard University, Cambridge, Massachusetts
(Received April 28, 1952)

Complete and quantitatively accurate solutions of the properties of electrically short transmitting and receiving antennas (of length $2h$ such that $2\pi h/\lambda_0 \leqq 1$) are obtained by determining the distributions of current that actually satisfy the integral equations. Components of current in phase and in phase-quadrature with the driving voltage or the incident electric field are evaluated together with the impedance, the effective length, and the gain. It is shown that when the King-Middleton method of solving Hallén's integral equation by iteration is applied correctly, quite accurate results are obtained even in a first-order solution. The greatest error in the first-order formulas is shown to be in the resistance, a small quantity of higher order compared with the reactance. The newly determined values are combined with the King-Middleton second-order results to obtain more complete and more accurate impedances and effective lengths in the range of $0 \leqq 2\pi h/\lambda \leqq 1.4$.

I. INTRODUCTION

A CENTER-DRIVEN cylindrical antenna of half-length h and radius a is electrically short by definition if its electrical half-length is less than unity. That is,

$$\beta_0 h = (2\pi h/\lambda_0) < 1. \tag{1}$$

This may be interpreted quantitatively as $\beta_0 h \leqq 0.5$, although the less severe condition, $\beta_0 h \leqq 1$, often is sufficiently restrictive. Most long-wave transmitting and receiving antennas are electrically short.

Although the properties of short antennas have been determined approximately by *assuming* the distribution of current to be triangular and the distribution of charge uniform, accurate results can not be derived on this basis. The short antenna is, of course, included in the general analysis of cylindrical antennas based on the integral equation of Hallén.[1–3] However, since the resistance of a short antenna is extremely small compared with the reactance, the usual method of solving the integral equation by iteration, using a real expansion parameter defined in terms of a real trial distribution of current (e.g., the zeroth-order current[2]) is not entirely satisfactory. In particular, it leads to results that are relatively much less accurate for the resistance than for the reactance when, owing to the complexity of higher order integrals, only one or two iterations can be evaluated. This situation can be improved by introducing as a trial function in the integral equation a complex distribution of current instead of a real distribution. It can be improved further for the electrically short antenna by using the ratio of vector-potential *difference* (referred to the end of the antenna) to current to define the expansion parameter rather than the ratio of vector potential to current since the former is much more nearly constant over the antenna than is the latter when the antenna is short. If these changes are made, two simultaneous integral equations, each involving only real quantities, are obtained, respectively, from the real and imaginary parts of the integral equation for the vector-potential difference. These equations may be arranged so that the one involves only the conductance, the other only the susceptance. However, in the general case, both components of current occur in each equation. It is to be expected that the separate solution of these two real equations by a method of iteration should lead to even more accurate results with a given number of iterations, especially for the conductance, than the conventional procedure of iterating in a single complex integral equation. This should be true particularly for very short and very long antennas. A study of the latter is reserved for another paper; the former are considered in the following.

Although the analysis of the short antenna may be obtained as a special case of the general one as outlined above, it can be carried out much more simply, directly, and accurately as a separate problem, using a procedure that is not applicable for longer antennas. This is a consequence of the fact that in the short antenna the component of current in phase with the driving voltage or incident field is so small that its contribution to the equation for the susceptance may be neglected. Hence, both the quadrature current and the susceptance may be determined and used in the solution of the equation for the conductance and the in-phase component of current.

II. INTEGRAL EQUATIONS FOR THE CENTER-DRIVEN ANTENNA

The integral equation for the thin cylindrical antenna of half-length h and radius a when center-driven by a discontinuity in scalar potential V_0 may be expressed as follows in terms of the axial component A_z of the vector potential on the surface of the antenna:[1–4]

* The research reported in this document was made possible through support extended Cruft Laboratory, Harvard University, jointly by the Navy Department (ONR), the Signal Corps of the U. S. Army, and the U. S. Air Force, under ONR Contract N5ori-76, T. O. 1.

[1] E. Hallén, Nova Acta Royal Soc. Sciences, Upsala **11**, 1 (1938).
[2] R. King and D. Middleton, Quart. Appl. Math. **3**, 302 (1946).
[3] D. Middleton and R. King, J. Appl. Phys. **17**, 273 (1946).
[4] R. W. P. King, *Theory of Linear Antennas* [Harvard University Press (to be published)].

$$4\pi\nu_0 A_z(z) = \int_0^1 I(\zeta')\left[\frac{\exp(-j\beta_0 \mathrm{hr}_1)}{r_1}+\frac{\exp(-j\beta_0 \mathrm{hr}_2)}{r_2}\right]d\zeta'$$

$$=\frac{-j4\pi}{\zeta_0}[C\cos\beta_0 h\zeta+\tfrac{1}{2}V_0\sin\beta_0 h\zeta], \quad (2)$$

where $\zeta=z/h$ is the normalized distance along the axis of the antenna, $r_1=[(\zeta'-\zeta)^2+\alpha^2]^{\frac{1}{2}}$ is the distance from the element $d\zeta'$ on the axis of the upper half of the antenna to the point $\zeta=z/h$ on the surface of the antenna, $r_2=[(\zeta'+\zeta)^2+\alpha^2]^{\frac{1}{2}}$ is the distance from the corresponding element $d\zeta'$ in the lower half of the antenna to the same point ζ; $\alpha=a/h$ is the normalized radius; C is an arbitrary constant of integration; $\zeta_0=120\pi$ ohms, $\nu_0=1/\mu_0=10^7/4\pi$ m/henry. The total axial current $I(\zeta')$ at $\zeta'=z'/h$ is complex. It may be separated into its real and imaginary parts and expressed in terms of the current $I(0)=V_0(G_0+jB_0)$ at the driving point as follows:

$$I(\zeta')=I''(\zeta')+jI'(\zeta')=[f_r(\zeta')G_0+jf_i(\zeta')B_0]V_0. \quad (3)$$

G_0 and B_0 are the input conductance and susceptance of the antenna; $f_r(\zeta')$ and $f_i(\zeta')$ are the real distribution functions for the components of current in phase and in phase-quadrature with the driving voltage.

The integral for the vector potential in Eq. (2) may be separated into four terms using Eq. (3):

$$4\pi\nu_0 A_z(z)=V_0G_0[\psi_r''(\zeta)+j\psi_r'(\zeta)]+V_0B_0[\psi''(\zeta)+j\psi_i'(\zeta)], \quad (4)$$

where

$$\psi_r''(\zeta)=\int_0^1 f_r(\zeta')\left[\frac{\cos\beta_0 \mathrm{hr}_1}{r_1}+\frac{\cos\beta_0 \mathrm{hr}_2}{r_2}\right]d\zeta', \quad (5a)$$

$$\psi_r'(\zeta)=-\int_0^1 f_r(\zeta')\left[\frac{\sin\beta_0 \mathrm{hr}_1}{r_1}+\frac{\sin\beta_0 \mathrm{hr}_2}{r_2}\right]d\zeta', \quad (5b)$$

$$\psi_i''(\zeta)=\int_0^1 f_i(\zeta')\left[\frac{\cos\beta_0 \mathrm{hr}_1}{r_1}+\frac{\cos\beta_0 \mathrm{hr}_2}{r_2}\right]d\zeta', \quad (5c)$$

$$\psi_i'(\zeta)=-\int_0^1 f_i(\zeta')\left[\frac{\sin\beta_0 \mathrm{hr}_1}{r_1}+\frac{\sin\beta_0 \mathrm{hr}_2}{r_2}\right]d\zeta'. \quad (5d)$$

Note that all quantities in (5a) through (5d) are real. The substitution of (4) and (5a), (5b), (5c), (5d) into (2) gives

$$\frac{4\pi\nu_0 A_z(z)}{V_0}=[G_0\psi_r''(\zeta)-B_0\psi_i'(\zeta)]+j[G_0\psi_r'(\zeta)+B_0\psi_i''(\zeta)]$$

$$=\frac{-j2\pi}{\zeta_0}[C'\cos\beta_0 h\zeta+\sin\beta_0 h\zeta], \quad (6a)$$

where $C'=2C/V_0$. This equation is valid for all values of ζ between 0 and 1. Let the vector-potential difference per unit voltage, *viz.*, $(4\pi\nu_0/V_0)[A_z(z)-A_z(h)]$, be introduced using (6a). Also let the following real difference functions be defined:

$$W_r\equiv\psi_r(\zeta)-\psi_r(1), \quad (7a)$$

$$W_i\equiv\psi_i(\zeta)-\psi_i(1). \quad (7b)$$

With this notation and the value of C' given by (6a) with $\zeta=0$, the following equations are obtained after separating the real and imaginary parts:

$$G_0W_r''-B_0W_i'=[G_0\psi_r''(0)-B_0\psi_i'(0)]\times[\cos\beta_0 h\zeta-\cos\beta_0 h], \quad (8a)$$

$$G_0W_r'+B_0W_i''=[G_0\psi_n'(0)+B_0\psi_i''(0)]\times[\cos\beta_0 h\zeta-\cos\beta_0 h]-\frac{2\pi}{\zeta_0}[\sin\beta_0 h\zeta-\sin\beta_0 h]. \quad (8b)$$

These are general expressions that apply to an antenna of any length. They may be rearranged so that one equation involves B_0, the other G_0.

The electrically short antenna presents a special problem owing to the fact that its resistance is very small compared with its reactance. In anticipation of the final results and to clarify the problem of retaining only the required terms, let it be stated that, subject to the general condition stated in Eq. (1), the following orders of magnitude are true:

$$B_0=\frac{1}{X_0}\sim\beta_0 h; \quad G_0=\frac{R_0}{X_0^2}\sim\beta_0^4 h^4. \quad (9)$$

It is evident, therefore, that it is possible to retain the principal term of order $\beta_0 h$ *and the first correction term* of order $\beta_0^3 h^3$ in B_0 in (8b) while all terms with G_0 as a factor are neglected.

Subject to Eq. (9) the general integrals in (5a–d) may be expanded as follows:

$$\psi''(\zeta)=\psi_0''(\zeta)+\psi_2''(\zeta)+\psi_4''(\zeta)+\cdots, \quad (10a)$$

$$\psi'(\zeta)=\psi_1'(\zeta)+\psi_3'(\zeta)+\psi_5'(\zeta)+\cdots, \quad (10b)$$

where

$$\psi_0''(\zeta)=\int_0^1 f(\zeta')\left(\frac{1}{r_1}+\frac{1}{r_2}\right)d\zeta; \quad (11a)$$

$$\psi_2''(\zeta)=\frac{-\beta_0^2h^2}{2}\int_0^1 f(\zeta')(r_1+r_2)d\zeta'; \quad (11b)$$

$$\psi_4''(\zeta)=\frac{\beta_0^4h^4}{24}\int_0^1 f(\zeta')(r_1^3+r_2^3)d\zeta'; \quad (11c)$$

$$\psi_1'(\zeta)=-2\beta_0 h\int_0^1 f(\zeta')d\zeta'; \quad (11d)$$

$$\psi_3'(\zeta)=\frac{\beta_0^3h^3}{6}\int_0^1 f(\zeta')(r_1^2+r_2^2)d\zeta'; \qquad (11e)$$

$$\psi_5'(\zeta)=\frac{-\beta_0^5h^5}{120}\int_0^1 f(\zeta')(r_1^4+r_2^4)d\zeta'. \qquad (11f)$$

Identifying subscripts r and t may be added to the f's and ψ's as required.

III. THE REACTANCE OF THE ELECTRICALLY SHORT ANTENNA

If only terms with orders of magnitude $\beta_0^3h^3$ and greater are retained, (8b) becomes independent of G_0 and may be solved for $X_0=-1/B_0$. Note that W_{t2}'' is of order $\beta_0^2h^2$.

$$-X_0=\frac{1}{B_0}=\frac{\zeta_0}{2\pi\beta_0h}\left[\frac{W_{t0}''}{1-\zeta}+\frac{\beta_0^2h^2W_{t0}''(1-\zeta^3)}{6(1-\zeta)^2}+\frac{W_{t2}''}{1-\zeta}-\frac{\beta_0^2h^2\psi_{t0}''(0)(1-\zeta^2)}{2(1-\zeta)}\right]. \qquad (12a)$$

For sufficiently small values of β_0h, the leading term adequately represents the entire solution. It is

$$-X_0=\frac{1}{B_0}=\frac{\zeta_0}{2\pi\beta_0h}\left[\frac{W_{t0}''}{1-\zeta}\right]. \qquad (12b)$$

This equation indicates that the quantity, $W_{t0}''=\psi_{t0}''(\zeta)-\psi_{t0}''(1)$, which is proportional to the principal part of the vector-potential difference, *varies linearly* with $\zeta=z/h$ along the antenna. Evidently, the solution of (12b) is at hand if the distribution function $f_t(\zeta')$, which makes W_t'' a linear function of ζ, can be determined. More conveniently, the required distribution function must satisfy $\psi_{t0}''(\zeta)$ in the relation

$$\frac{W_{t0}''}{1-\zeta}=\frac{\psi_{t0}''(\zeta)-\psi_{t0}''(1)}{1-\zeta}=\text{const}=K_{t0}'', \qquad (13)$$

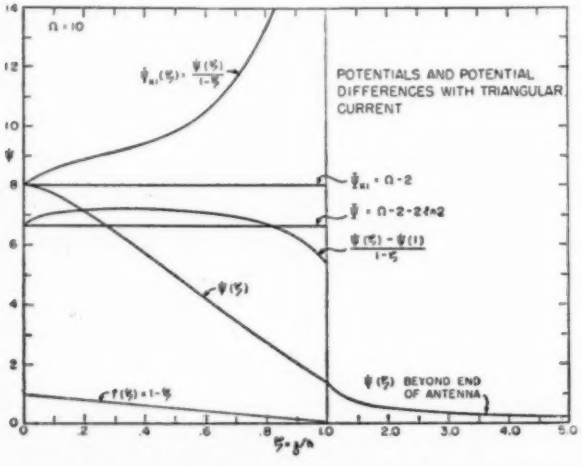


FIG. 1. Potential and potential differences with triangular current; $\psi(\zeta)$ stands for $\psi_{t0}''(\zeta)$.

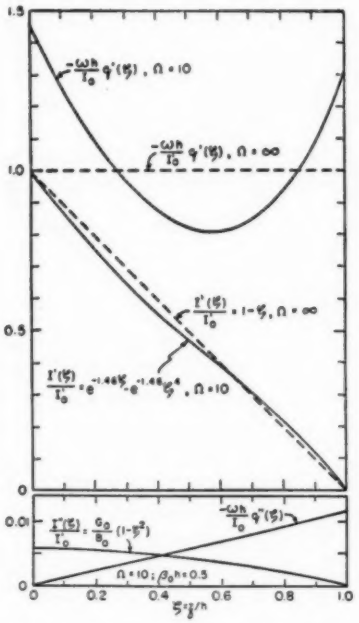


FIG. 2. Distributions of current and charge on short antennas with $\Omega=2\ln(2h/a)=10$ and ∞.

for $0\leqq\zeta\leqq1$. Subject to (11) and (9), the leading term $\psi_{t0}''(\zeta)$ in $\psi_t''(\zeta)$ as defined in (4c) is

$$\psi_{t0}''(\zeta)\doteq\int_0^1 f_t(\zeta')\left[\frac{1}{r_1}+\frac{1}{r_2}\right]d\zeta'. \qquad (14)$$

It is verified readily that the linear distribution

$$f(\zeta')=1-\zeta', \qquad (15a)$$

when substituted in (14) for use in (13), does not give a constant for antennas with a/h finite. The function of ζ' defined by (13) with the distribution (15a) is shown in Fig. 1 together with the linear distribution (15a) for $\Omega=2\ln(2h/a)=10$. The function $\psi_{t0}''(\zeta)$ also is shown. An inspection of these curves discloses at once that a distribution of current $f(\zeta')$, which has relatively greater amplitudes than $(1-\zeta')$ between $\zeta'=0$ and 0.1 and between $\zeta'=0.8$ and 1.0, is required to make the ratio (13) constant. Such a distribution may be constructed by combining the linear distribution (15a) with a number of other linear distributions with appropriate amplitudes but extended only over *small fractions* of the antenna. Since a distribution consisting of a superposition of a large number of linear distributions over parts of the antenna is awkward in the subsequent evaluation of higher-order terms, an essentially equivalent distribution has been constructed in the following analytically simple form:

$$I'(\zeta')/I_0'=f_t(\zeta')=\exp(-\delta\zeta')-e^{-\delta}\zeta'^n. \qquad (15b)$$

This distribution is represented in Fig. 2 for $\Omega=10$, for which the values, $\delta=1.46$, $n=4$, make (15b) an excellent approximation of the previously determined, analyti-

cally much more complicated current. If (13) is evaluated from (14) using (15b), it is found to be essentially constant for all values of ζ. This is illustrated in Fig. 3 where the distribution (15b) is shown together with (13) and (14) evaluated using this distribution. Note that the principal component of current on the antenna $I'(\zeta')/I_0'=f(\zeta')$ has a distribution that departs considerably from a linear form. This is a necessary consequence of the fact that the vector potential itself is linear. Significant are the rapid drop in current as ζ is increased from zero and the increase in current near the end at $\zeta=1$. It is seen in Fig. 3 that the vector potential, which is proportional to $\psi_{t0}''(\zeta)$, due to the exponential distribution (15b) is linear except very near $\zeta=0$. The requirement that the vector potential must be linear down to $\zeta=0$ is a consequence of the assumed driving condition consisting of a discontinuity in scalar potential at $\zeta=0$. Since in practice this idealized generator is approximated by a transmission line in the limit as the spacing of the line approaches but never reaches zero, the vector potential in any practical case is rounded off near $\zeta=0$. It follows that, whereas the distribution (15b) with $\delta=1.46$ does not satisfy the idealized integral equation very near $\zeta=0$, it does agree with it if the halves of the antenna are connected to a transmission line and separated a very small distance 2δ. Evidently, account must be taken of terminal-zone effects as described in the literature.[4-6] The idealized but physically impossible distribution function evidently must be one differing from (15) only near $\zeta=0$ sufficiently to make $\psi_{t0}(0)=8$. Excluding values very near $\zeta=0$ and $\zeta=1$, the average constant value K_t'' of (12) is

$$K_{t0}''=6.60. \tag{16}$$

Except for the point at $\zeta=0$, no part of the function (12) using (15) differs from 6.60 by as much as one percent. It follows that the modified exponential distribution (15) is the *true distribution* of the component of current I_s' in phase quadrature with the driving voltage for an antenna with $\Omega=10$.

With (16) and $\zeta_0=120\pi$ ohms it follows from (11) that the limiting formula for X_0 as $\beta_0 h$ is made smaller is

$$\Omega=10:\quad X_0=-\frac{1}{B_0}=-\frac{6.60\times 60}{\beta_0 h}=-\frac{396.0}{\beta_0 h}\ \text{ohms}. \tag{17}$$

The variation of X_0 with $\beta_0 h$ as $\beta_0 h$ is increased from near zero to 0.5 or even 1.0 may be determined more accurately by solving (12a) instead of (12b). Strictly, this requires a redetermination of the distribution of current. However, since it is a question merely of determining a higher-order correction term it is adequate to assume the same distribution of current so that the first term in the brackets in (12a) is simply $K_t''=6.60$ as given by (16). Since for each value of $\beta_0 h$ the sum of the last three terms in (12a) must be essentially a constant independent of ζ, the value at $\zeta=0$ may be assumed to be a satisfactory approximation of this constant. With $\delta=1.46$, for $\Omega=10$,

$$W_{t2}''(0)=\psi_{t2}''(0)-\psi_{t2}''(1)=0.367\beta_0^2h^2. \tag{18a}$$

From Fig. 3 it follows that

$$\psi_{t0}''(0)=8. \tag{18b}$$

Substitution of these several values in (12a) with $\zeta=0$ gives the following expression for the reactance of the electrically short antenna with $\Omega=10$:

$$-X_0=\frac{1}{B_0}=\frac{396.0}{\beta_0 h}[1-0.383\beta_0^2h^2]\ \text{ohms};\quad \beta_0 h\lesssim 1. \tag{19}$$

Points computed from (19) are shown in Fig. 4 on the curve for $-X_0$ with $\Omega=10$.

IV. THE RESISTANCE OF THE ELECTRICALLY SHORT ANTENNA

With the susceptance determined from (8b) the conductance and its first correction term may be evaluated from (8a). Subject to (9) and using (11) and the series expansion of the cosine function, (8a) may be expressed as in (20a). Note that since G_0 is of the order of magnitude of $\beta_0^4h^4$ and B_0 of magnitude $\beta_0 h$, it is necessary to include terms up to order $\beta_0^6h^6$. Since $\psi_1'(\zeta)$ actually is independent of ζ as seen in (11), $W_1'=0$, so that the lowest-order terms in (20a) are $\beta_0^4h^4$. The rearranged

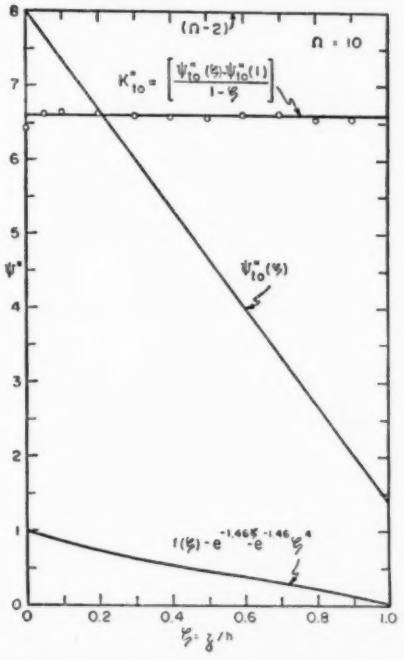


FIG. 3. The function $I_{t0}''(\zeta)$ and the constant K_{t0}'' for a short antenna with a modified exponential distribution of current. The small circles denote computed points.

[4] R. King, J. Appl. Phys. **20**, 832 (1949).
[5] R. King and K. Tomiyasu, Proc. Inst. Radio Engrs. **37**, 1134 (1949).

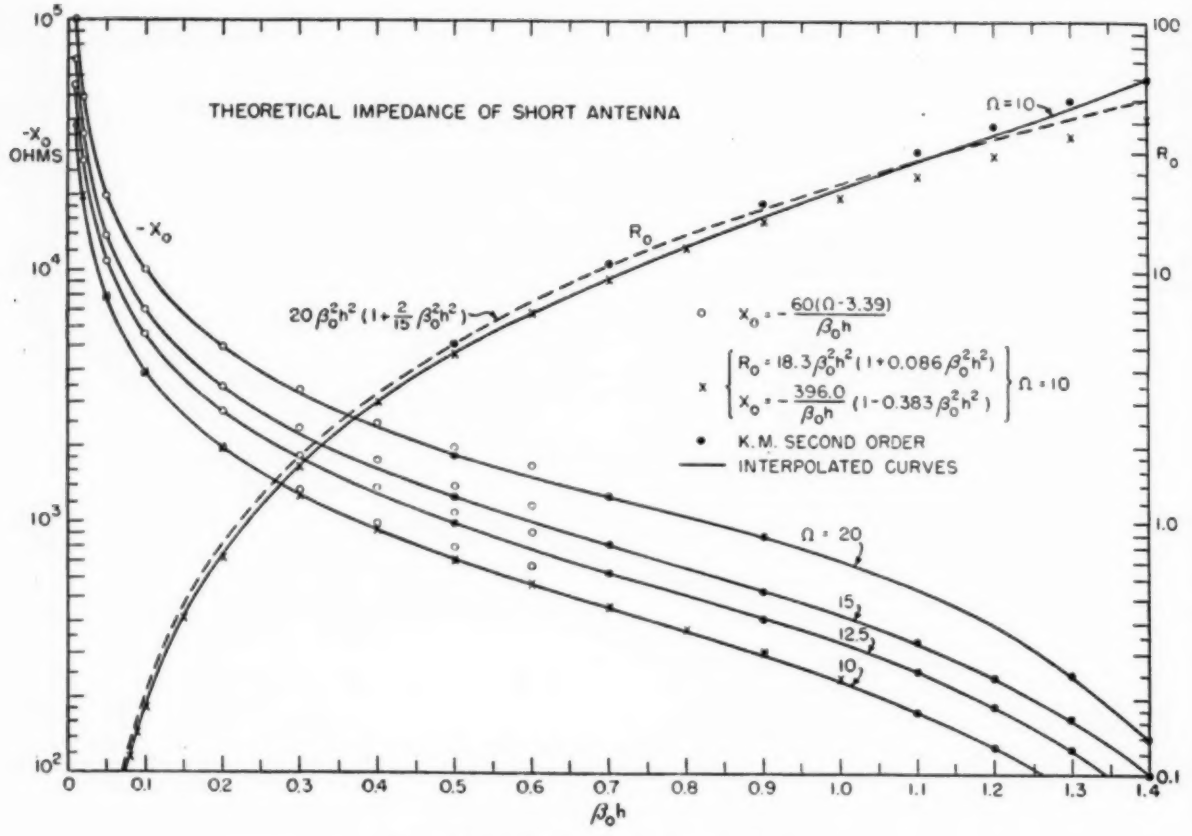


FIG. 4. Theoretical impedance of a short antenna.

equation is

$$G_0W_{r0}''+G_0[W_{r2}''-\tfrac{1}{2}\beta_0^2h^2\psi_{r0}''(0)(1-\zeta^2)]$$
$$=B_0[W_{t3}'-\tfrac{1}{2}\beta_0^2h^2\psi_{t1}'(0)(1-\zeta^2)]$$
$$+B_0\left[W_{t5}'-\frac{\beta_0^2h^2}{2}\psi_{t3}'(0)(1-\zeta^2)\right.$$
$$\left.+\frac{\beta_0^4h^4}{24}\psi_{t1}'(0)(1-\zeta^4)\right]. \quad (20)$$

The leading terms of order $\beta_0^4h^4$ are:

$$G_0\left[\frac{W_{r0}''}{1-\zeta^2}\right]=B_0\left[\left(\frac{W_{t3}'}{1-\zeta^2}\right)-\tfrac{1}{2}\beta_0^2h^2\psi_{t1}'(0)\right]. \quad (20b)$$

Since the distribution of current $f_t(\zeta')$, which determines $\psi_t'(\zeta)$ and $W_t=W_t''+jW_t'$, already has been obtained (and is given by (15b) with $n=4$ and $\delta=1.46$ for $\Omega=10$) the quantities on the right in (20b) may be evaluated. The substitution of (15b) with $n=4$ in $\psi_{t1}'(\zeta)$ and $\psi_{t3}'(\zeta)$ as given by (11) leads to

$$\frac{W_{t3}'}{1-\zeta^2}=\frac{-\beta_0^3h^3}{3}\left[\frac{1-e^{-\delta}}{\delta}-\frac{e^{-\delta}}{5}\right]=K_{t3}'. \quad (21)$$

With $\delta=1.46$ for $\Omega=10$, $\psi_{t1}'(0)=-0.9589\beta_0h$, $\psi_{t3}'(0)=0.02778\beta_0^3h^3$, $K_{t3}'=-0.1598\beta_0^3h^3$.

Since the entire right side of (20b) is equal to a constant, it follows that

$$\frac{W_{r0}''}{1-\zeta^2}=\frac{\psi_{r0}''(\zeta)-\psi_{r0}''(1)}{1-\zeta^2}=\text{const}=K_{r0}''. \quad (22)$$

Evidently, the evaluation of K_{r0}'' involves the simultaneous determination of the distribution function $f_r(\zeta')$ of the small component of current in phase with the driving voltage. The correct value of $f_r(\zeta')$ is that which reduces the function on the left in (20) to a constant for all values of ζ in the range 0 to 1. The integral involved is (11a) with subscript r on ψ and f.

It is shown in Fig. 5 that the parabolic distribution, $f(\zeta')=1-\zeta'^2$, when substituted in (11a) does not yield a constant when used in (18). On the other hand, by modifying the parabolic distribution by adding to it the distribution $p(\zeta)$ shown at the bottom in Fig. 5 to obtain the distribution plotted in Fig. 5 as $f_r(\zeta)=1-\zeta^2+p(\zeta)$, and then substituting this function in (11a) for use in (25), the following are obtained with $\Omega=10$: $K_{r0}''=6.90$; $\psi_{r0}''(0)=9.0$. It follows that the curve $f_r(\zeta)=1-\zeta^2+p(\zeta)$ in Fig. 5 is the *correct distribution* for the component of current in phase with the driving voltage.

The substitution of the several numerical values for $\Omega=10$ in (20b) gives:

$$\frac{G_0}{B_0}=-\frac{R_0}{X_0}=\left[\frac{K_{t3}-\frac{1}{2}\psi_{t1}'(0)\beta_0^2h^2}{K_{r0}''}\right]$$

$$=\left[\frac{-0.1598+0.4794}{6.90}\right]=0.0463\beta_0^3h^3. \quad (23)$$

With (17), it follows that for a sufficiently short antenna with $\Omega=10$, the input resistance is

$$R_0=18.34\beta_0^2h^2; \quad (\beta_0h<0.5). \quad (24)$$

In order to obtain the first-order correction for R_0 it is necessary to solve (20a) instead of (20b). If it is assumed that the distribution of current is essentially the same as that determined from the leading terms in (20b), and the correction terms are evaluated at $\zeta=0$, (20a) may be expressed as follows:

$$\frac{G_0}{B_0}=\frac{-R_0}{X_0}=\frac{K_{t3}'-\frac{1}{2}\psi_{t1}'(0)\beta^2h^2+W_{t5}'(0)-\frac{1}{2}\psi_{t3}'(0)\beta^2h^2+\frac{1}{24}\psi_{t1}'(0)\beta_0^4h^4}{K_{r0}''+W_{r2}''(0)-\frac{1}{2}\beta_0^2h^2\psi_{r0}''(0)}. \quad (25)$$

The numerical value of this ratio for $\Omega=10$ is

$$\frac{G_0}{B_0}=\frac{-R_0}{X_0}=0.0463\beta_0^3h^3(1+0.469\beta_0^2h^2), \quad (26)$$

so that with (19), the resistance of the short antenna with $\Omega=10$ is

$$R_0=18.3\beta_0^2h^2(1+0.086\beta^2h^2); \quad \beta_0h\leqq0.9. \quad (27)$$

This function is shown in Fig. 4. The following quantities not previously determined are involved in the evaluation of (25) for

$$\Omega=10: \quad \psi_{t5}'(0)=-0.000587\beta_0^5h^5;$$
$$\psi_{t5}'(1)=-0.0169\beta_0^5h^5;$$
$$W_{t5}'(0)=\psi_{t5}'(0)-\psi_{t5}'(1)=0.0163\beta_0^5h^5;$$
$$\psi_{r}^{2\prime\prime}(0)=-\beta_0^2h^2/4;$$
$$\psi_{r2}''(1)=-2\beta_0^2h^2/3;$$
$$W_{r2}''(0)=\psi_{r2}''(0)-\psi_{r2}''(1)=5\beta_0^2h^2/12.$$

The final expression for the impedance of the short antenna with $\Omega=10$ is

$$Z_0=18.3\beta_0^2h^2(1+0.086\beta_0^2h^2)$$
$$-j\frac{396.0}{\beta_0h}(1-0.383\beta_0^2h^2). \quad (28)$$

This is accurate for $\beta_0h\leqq0.5$ and an excellent approximation for $\beta_0h\leqq0.9$. Corresponding formulas for other values of Ω may be determined in the same manner. A comparison of this formula with the corresponding impedance obtained by solving the integral Eq. (2) by iteration following King and Middleton is carried out in Sec. VII.

An alternative and equivalent method for determining the resistance R_0 in (27) is by applying the Poynting-vector theorem to the radiation field computed using the distribution of current in (15b). Actual evaluation for $\Omega=10$ yields (27), as it should. The corresponding value for $\Omega=\infty$ and a triangularly distributed current is

$$R_0=20\beta_0^2h^2\left(1+\frac{2}{15}\beta_0^2h^2\right). \quad (29)$$

V. DISTRIBUTIONS OF CURRENT AND CHARGE ON THE SHORT CENTER-DRIVEN ANTENNA

Since the evaluation of the impedance (28) of the electrically short antenna with $\Omega=10$ involves the simultaneous determination of the distribution of current, this latter may be written down directly. With (3) it has the following form for $\Omega=10$:

$$I(\zeta)=I''(\zeta)+jI'(\zeta), \quad (30)$$

where

$$I''(\zeta)=I_0''f_r(\zeta)=I_0''[1-\zeta^2+p(\zeta)]\doteq I_0''[1-\zeta^2];$$
$$I_0''=G_0V_0=I_0'\frac{G_0}{B_0}; \quad (31)$$

$$I'(\zeta)=I_0'f_t(\zeta)=I_0'[\exp(-\delta\zeta)-\epsilon^{-\delta}\zeta^n]; \; I_0'=B_0V_0; \quad (32)$$

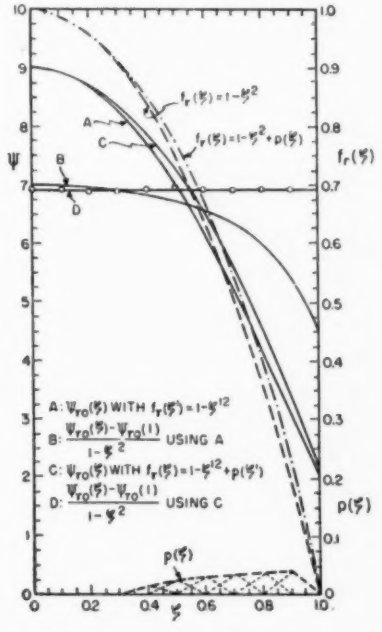


FIG. 5. Potentials and potential differences related to the component of current in phase with the driving voltage in a short center driven antenna and in phase quadrature with the incident electric field in a short receiving antenna.

where for $\Omega=10$, $p(\zeta)$ is in Fig. 5, $\delta=1.46$ and $n=4$. Since $p(\zeta)$ is a very small term in a very small component of current, it may be neglected in most applications. For the infinitely thin antenna ($\Omega=\infty$),

$$I''(\zeta)=0; \quad I'(\zeta)=I_0'(1-\zeta). \tag{33}$$

These distributions are shown in Fig. 2.

The distribution of charge is derived from the equation of continuity, $dI/dz+j\omega q=0$. Thus, neglecting $p(\zeta)$,

$$q=j(q''+jq'), \tag{34}$$

where

$$q''=\frac{1}{\omega h}\frac{dI''}{d\zeta}=-\frac{2I_0''}{\omega h}\zeta=-\frac{2I_0'G_0}{\omega hB_0}\zeta, \tag{35a}$$

$$q'=\frac{1}{\omega h}\frac{dI'}{d\zeta}=-\frac{I_0\delta}{\omega h}\left(\exp-\delta\zeta+\frac{ne^{-\delta}}{\delta}\zeta^{n-1}\right). \tag{35b}$$

For the infinitely thin antenna

$$q''=0; \quad q'=-\frac{I_0'}{\omega h}. \tag{36}$$

The quantities $-(\omega h/I_0')q'(\zeta)$ and $-(\omega h/I_0')q''(\zeta)$ are shown in Fig. 2 for $\Omega=10$ for which $\delta=1.46$ and $n=4$; $-(\omega h/I_0')q'(\zeta)=1$ is shown for $\Omega=\infty$.

VI. THE ELECTRICALLY SHORT RECEIVING ANTENNA

The integral equation for the electrically short receiving antenna is[1, 4, 7]

$$4\pi\nu_0 A_z(z)=\int_0^1 I(\zeta')\left[\frac{\exp-j\beta_0\,\mathrm{hr}_1}{r_1}+\frac{\exp-j\beta_0\,\mathrm{hr}_2}{r_2}\right]d\zeta'$$

$$=\frac{-j4\pi}{\zeta_0}[C\cos\beta_0h\zeta+U], \tag{37}$$

where the notation is the same as in (2). When the receiving antenna is parallel to the incident electric field E at the center of the antenna, as is assumed here for simplicity,

$$U=E/\beta_0. \tag{38}$$

Let

$$I(\zeta')=I''(\zeta)+jI'(\zeta)=I_0''f_r(\zeta)+jI_0'f_s(\zeta), \tag{39}$$

where $I''(\zeta)$ is the component of current in phase, $I'(\zeta)$ the component in phase-quadrature with the incident electric field, and I_0'' and I_0' are the components of current at the center, $\zeta=0$, of the antenna. With this notation, (37) becomes

$$4\pi\nu_0A_z(z)=I_0''[\psi_s''(\zeta)+j\psi_s'(\zeta)]+I_0'[\psi_r'(\zeta)+j\psi_r'(\zeta)], \tag{40}$$

where $\psi_r''(\zeta)$ and $\psi_r'(\zeta)$ are defined in (4a, b), $\psi_s''(\zeta)$ and $\psi_s'(\zeta)$ by (4c, d) with a subscript s replacing t.

[7] R. King and C. W. Harrison, Jr., Proc. Inst. Roy. Engrs. **32**, 18, 35 (1944).

If phase is referred to U, (37) may be expressed as follows:

$$\frac{4\pi\nu_0A_z(z)}{U}=[u_0''\psi_s''(\zeta)-u_0'\psi_r'(\zeta)]+j[u_0''\psi_s'(\zeta)+u_0'\psi_r''(\zeta)]$$

$$=\frac{-j4\pi}{\zeta_0}[C'\cos\beta_0h\zeta+1], \tag{41}$$

where $C'=C/U$, and $u_0''=I_0''/U$, $u_0'=I_0'/U$ are the normalized components of current. Now let the vector-potential differences be introduced in (41) as in (7a, b) but with the subscript s replacing t

$$[u_0''W_s''(\zeta)-u_0'W_r'(\zeta)]+j[u_0''W_s'(\zeta)+u_0'W_r''(\zeta)]=\frac{-j4\pi}{\zeta_0}C'[\cos\beta_0h\zeta-\cos\beta_0h]. \tag{42}$$

It is now convenient to substitute the value of C' from (41) with $\zeta=0$ into (42). The real and imaginary parts of the result are

$$[u_0''W_s''(\zeta)-u_0'W_r'(\zeta)]=[u_0''\psi_s''(0)-u_0'\psi_r'(0)]\times[\cos\beta_0h\zeta-\cos\beta_0h]; \tag{43a}$$

$$[u_0''W_s'(\zeta)+u_0'W_r''(\zeta)]=\left[u_0''\psi_s'(0)+u_0'\psi_r''(0)+\frac{4\pi}{\zeta_0}\right]\times[\cos\beta_0h\zeta-\cos\beta_0h]. \tag{43b}$$

These equations may be solved formally for u_0'' and u_0'.

For the short antenna,

$$\cos\beta_0h\zeta-\cos\beta_0h=\tfrac{1}{2}\beta_0^2h^2(1-\zeta^2)\left[1-\frac{1}{12}\beta_0^2h^2\left(\frac{1-\zeta^4}{1-\zeta^2}\right)\right], \tag{44}$$

$$W''(\zeta)=W_0''(\zeta)+W_2''(\zeta)+\cdots, \tag{45a}$$

$$\psi''(\zeta)=\psi_0''(\zeta)+\psi_2''(\zeta)+\cdots, \tag{45b}$$

$$W'(\zeta)=W_3'(\zeta)+W_5'(\zeta)+\cdots, \tag{45c}$$

$$\psi'(\zeta)=\psi_1'(\zeta)+\psi_3'(\zeta)+\cdots. \tag{45d}$$

Note that $W_1'(\zeta)=0$ since $\psi_1'(\zeta)$ is independent of ζ. It follows that $W'(\zeta)$ is of the order of $\beta_0^3h^3$ and higher, whereas $W''(\zeta)$ is of the order of unity. If terms of order not higher than $\beta_0^2h^2$ are to be retained in (43), *all* terms in $W_s'(\zeta)$ may be neglected; moreover, only the leading term in the trigonometric factor is required. Thus, (43b) reduces to

$$u_0'\left[\frac{W_{r0}''(\zeta)}{1-\zeta^2}+\frac{W_{r2}''(\zeta)}{1-\zeta^2}-\tfrac{1}{2}\psi_r''(0)\beta^2h^2\right]=\frac{2\pi}{\zeta_0}\beta_0^2h^2. \tag{46}$$

This includes a leading part of order $\beta_0^2h^2$ and a first-

order correction with factor $\beta_0^4h^4$. The leading term on the left gives

$$u_0'\left[\frac{W_{r0}''(\zeta)}{1-\zeta^2}\right]=\frac{2\pi}{\zeta_0}\beta_0^2h^2. \tag{47a}$$

But it has already been shown that with a modified parabolic distribution of current,

$$\frac{W_{r0}''(\zeta)}{1-\zeta^2}=\text{const}=K_{r0}''. \tag{47b}$$

By inserting $\zeta=0$ in the higher order terms, the following result is obtained:

$$u_0'[K_{r0}''+W_{r2}'(0)-\tfrac{1}{2}\beta^2h^2\psi_r''(0)]=\frac{2\pi}{\zeta_0}\beta_0^2h^2, \tag{48}$$

so that

$$u_0'=\frac{2\pi\beta_0^2h^2}{\zeta_0K_{r0}''}\left[1-\frac{W_{r2}''(0)}{K_{r0}''}+\frac{\frac{1}{2}\beta^2h^2\psi_r''(0)}{K_{r0}''}\right]. \tag{49}$$

For $\Omega=10$, $K_{r0}''=6.90$; $\psi_r''(0)=9$, $W_{r2}'(0)=0.4545\beta_0^2h^2$. Hence

$$u_0'=\frac{\beta_0^2h^2}{60\times6.90}[1+0.444\beta_0^2h^2]. \tag{50}$$

It is readily verified from (43a) that $I''(\zeta)$ is distributed essentially like $I'(\zeta)$ so that all subscripts s may be changed to r. It follows directly that u_0' is of the order of magnitude of $\beta_0^2h^2$, u_0'' of the order of magnitude of $\beta_0^5h^5$. Accordingly, u_0'' may be neglected in determining the effective length and the gain of the antenna. For this purpose the current in the antenna, with $\Omega=10$, is given by

$$I(\zeta)\doteq jI_0'(1-\zeta^2+p(\zeta))j\doteq I_0'(1-\zeta^2), \tag{51a}$$

where

$$I_0'=Uu_0'=0.00242U\beta_0^2h^2(1+0.444\beta^2h^2). \tag{51b}$$

Note that $I(\zeta)$ is in phase-quadrature with the incident electric field.

The effective electrical half-length β_0h_e of a symmetrical antenna of half-length h is defined as follows:[4,7]

$$\beta_0h_e=\beta_0(h_e''+jh_e')=\tfrac{1}{2}u_0(0)Z_0\doteq\frac{j}{2}u_0'(0)Z_0. \tag{52}$$

It is assumed that the electric field is parallel to the antenna. For $\Omega=10$, using (50) and (34) in (52), the effective electrical half-length is

$$\beta_0h_e''=0.47825\beta_0h(1+0.061\beta_0^2h^2)\doteq|\beta_0h_e|, \tag{53a}$$

$$\beta_0h_e'=0.00221\beta_0^4h^4(1+0.530\beta_0^2h^2)\doteq0. \tag{53b}$$

The function $h_e/\lambda_0\doteq h_e'/\lambda_0$ as evaluated from (53a) is plotted in Fig. 6. Note that effective length of electrically short antennas is longer for thin than for thick antennas. In particular, for $\beta_0h\leq1$, the curve for $\Omega=10$ in Fig. 6 lies below that for $\Omega=\infty$, which is given by

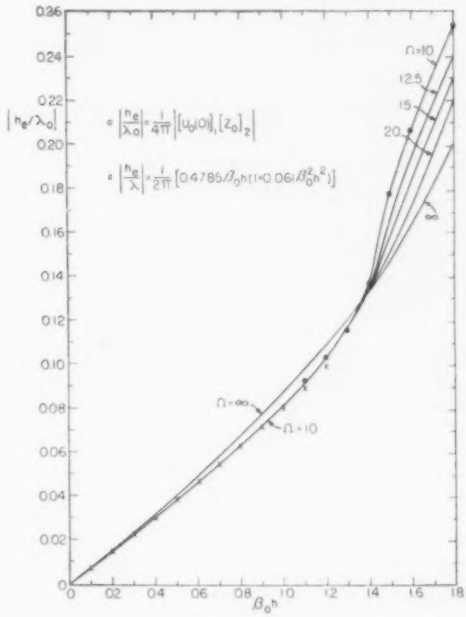


FIG. 6. The normalized magnitude of the theoretical effective length of a short receiving antenna.

$$\beta_0h_e=\tan(\beta_0h/2)\doteq\frac{\beta_0h}{2}\left(1+\frac{\beta_0^2h^2}{12}\right). \tag{54}$$

The important fact that all curves for effective lengths of antennas with $\Omega<\infty$ cross the curve for $\Omega=\infty$ near resonance is considered in Sec. VIII.

The directivity or gain of the receiving antenna parallel to the electric field ($\theta=\pi/2$) may be defined as follows:[4,7]

$$D=\frac{\zeta_0}{\pi}\frac{[\beta_0h_e]^2}{R_0}. \tag{55}$$

With (53a) and the real part of (28), this gives:

$$\Omega=10:\quad D=1.500(1+0.036\beta_0^2h^2). \tag{56}$$

Owing to the reciprocal theorem, the directivity or gain also can be defined in terms of the radiation field of the same antenna when driven. This is obtainable by conventional methods, but using the distribution of current (15b), not the commonly assumed triangular current. For $\Omega=10$, the result obtained in this manner is precisely (56), as it should be. The gain of the electrically short antenna is included in Fig. 7, where the range $\beta_0h\geq0.6$ is computed from (56) for $\Omega=10$. The corresponding formula for the infinitely thin antenna with a triangular current is given by (55) with (54) and (34b). It is

$$\Omega=\infty:\quad D(\theta)=\tfrac{3}{2}[1+\beta_0^2h^2/30]. \tag{57}$$

Evidently, the directivity of electrically short antennas is quite insensitive to thickness *provided this is suffi-*

ciently small. The very slightly greater rate of increase of the gain for $\Omega=10$ than for $\Omega=\infty$ apparently does not continue beyond $\beta_0 h=0.6$ owing to a somewhat more rapid increase of R_0 with $\beta_0 h$ than given by (34a). This is discussed in Sec. VIII. Actually, the gain or directivity decreases with increasing radius a. However, when $\beta_0 a=2\pi a/\lambda_0$ is very small compared with one, this decrease is of the order of magnitude of $\beta_0^2 a^2$, so that it is insignificant in (56). If $\beta_0 a$ is increased sufficiently, gains very considerably below 1.5 may be obtained.

VII. COMPARISON OF THE NEW ANALYSIS WITH THE RESULTS OF THE KING-MIDDLETON EXPANSION FOR THE VERY SHORT CENTER-DRIVEN ANTENNA

The integral equation for the center-driven antenna (2) was solved originally by Hallén using a method of iteration involving a series in reciprocal powers of the parameter $\Omega=2\ln(2h/a)$. Owing to the complexity of the integrals involved in higher order terms, King and Middleton modified Hallén's procedure by introducing a different expansion parameter which achieved somewhat more rapid convergence and hence greater accuracy in the first few terms. The King-Middleton method involved expressing the current $I(\zeta')$ at $\zeta'=z'/h$ in terms of the current $I(\zeta)$ at $\zeta=z/h$ using the distribution function $g(\zeta,\zeta')\doteq I(\zeta')/I'\zeta)$. Thus,

$$4\pi\nu_0 A_z(z)=I(\zeta)\Psi(\zeta)=I(\zeta)[\Psi+\gamma(\zeta)]$$

$$=\frac{-j4\pi}{\zeta_0}[C\cos\beta_0 h\zeta+\tfrac{1}{2}V_0\sin\beta_0 h\zeta], \quad (58)$$

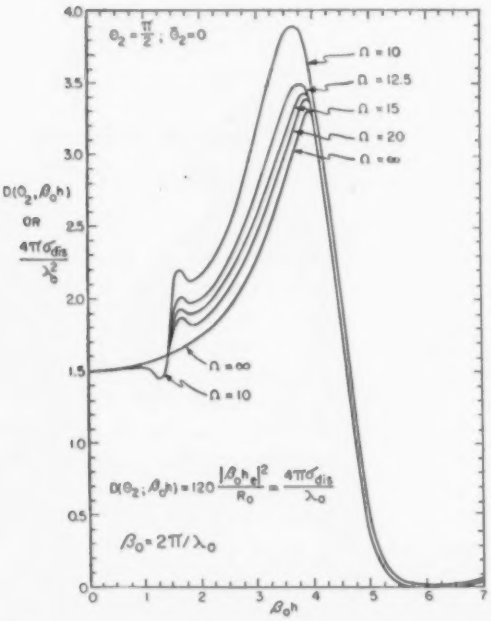


FIG. 7. The theoretical directivity or gain $D(\theta_2, \beta_0 h)$ of a cylindrical antenna.

where

$$\Psi(\zeta)\equiv\Psi+\gamma(\zeta)\equiv\int_0^1 g(\zeta,\zeta')\times\left[\frac{\exp(-j\beta_0 hr_1)}{r_1}+\frac{\exp(-j\beta_0 hr_2)}{r_2}\right]d\zeta'. \quad (59)$$

By an appropriate choice of the distribution function $g(\zeta,\zeta')$ the function $\Psi(\zeta)=4\pi\nu_0 A_z(z)/I(\zeta)$ is separated into a principal real part Ψ and a complex correction term $\gamma(\zeta)$ that is small except near the ends $\zeta=1$. The definition of a constant real function Ψ is largely arbitrary. It depends on the plausible assumption that the ratio of vector potential at the surface of a cylindrical conductor to the total current in the conductor is determined principally by elements of current adjacent to the point of calculation of the vector potential. In the original analysis[2,3] the validity of this assumption was verified specifically for $\beta_0 h=\pi/2$ and $\beta_0 h=\pi$ and the constant $\Psi=\Psi_{K1}$ was defined by

$$\Psi_{K1}=|\Psi(0)|/\sin\beta_0 h; \quad \beta_0 h\leqq\pi/2, \quad (60a)$$

$$\Psi_{K1}=|\Psi(h-\lambda_0/4)|; \quad \beta_0 h\geqq\pi/2, \quad (60b)$$

with $g(\zeta,\zeta')$ represented by the zeroth-order current in the form

$$g(\zeta,\zeta')=\sin\beta_0 h(1-\zeta')/\sin\beta_0 h(1-\zeta). \quad (60c)$$

It was verified also that no significant change in Ψ_{K1} as defined in (60a, b) was observed when an approximate first-order instead of zeroth-order current was introduced to define $g(\zeta,\zeta')$. Originally, no lower limit was specified in (60a) and it was assumed implicitly that $\Psi(\zeta)$ was sensibly constant at Ψ_{K1} as defined in (60a) in the limit as $\beta_0 h$ approached zero. It was shown that in this limit, Ψ_{K1} approaches $\Omega-2$. Evidently, the essence of the King-Middleton method is to find an approximate current distribution function $g(\zeta,\zeta')$ which leads to a function $\Psi(\zeta)$ that is sensibly constant over the antenna and thus permits the definition of this constant value as an appropriate expansion parameter. For electrically short antennas the component of current in phase with the driving voltage is so small compared with the quadrature current and the conductance is so small compared with the susceptance that an accurate determination of the conductance or resistance cannot be achieved in a first or even second-order solution with a real zeroth-order distribution function $g(\zeta,\zeta')$.

Using the zeroth-order distribution function and $\Psi_{K1}=\Omega-2$ as defined in (60a), the first-order impedance is given by $[Z_0]_1=[R_0]_1+j[X_0]_1$, where

$$[R_0]_1=20\beta_0^2 h^2\left[\frac{1+(2\ln 4/\Omega-2)}{1+(\ln 4/\Omega-2)^2}\right], \quad (61a)$$

$$[X_0]_1=\frac{-60}{\beta_0 h}\left[\frac{\Omega-2}{1+(\ln 4/\Omega-2)}\right]. \quad (61b)$$

For $\Omega=10$, the first-order values are

$$[R_0]_1=19.56\beta_0^2h^2 \text{ ohms};$$
$$[X_0]_1=-\frac{60\times 6.82}{\beta_0 h} \text{ ohms}. \quad (61c)$$

The correct values for $\Omega=10$ have been given in (34a). They are

$$R_0=18.3\beta_0^2h^2 \text{ ohms}; \quad X_0=-\frac{60\times 6.60}{\beta_0 h} \text{ ohms}. \quad (62)$$

Thus, the *first-order* King-Middleton values are in error for very short antennas by 6.9 percent in the resistance and 3.3 percent in the reactance.

In view of the fact that the ratio of vector potential to zeroth-order current as defined by $\Psi_{K1}(\zeta)=\Psi(\zeta)/(1-\zeta)$ is not at all constant near $\Omega-2$ as is seen from Fig. 1, the basic requirement of the King-Middleton method is violated. In view of this fact the degree of accuracy in the first-order solution is very satisfactory.

An alternative procedure for defining the expansion parameter (which better satisfies the King-Middleton requirement of reasonable constancy in the generating function $\Psi(\zeta)$ or its equivalent for short antennas) is in terms of the vector-potential difference $A_z(\zeta)-A_z(1)$ between the point of calculation at $\zeta=z/h$ along the conductor and the end of the conductor at $\zeta=z/h=1$. This difference is proportional to $\Psi(\zeta)-\Psi(1)$. It is evident from Fig. 1 that the ratio of this function to the zeroth-order current, i.e., $[\Psi(\zeta)-\Psi(1)]/(1-\zeta)$, is much more nearly constant at

$$\Psi_{Dt}\equiv\Psi(0)-\Psi(1)=\Omega-2-\ln 4 \quad (63)$$

than is $\Psi(\zeta)/(1-\zeta)$ as defined in (60a) at $\Psi_{K1}=\Omega-2$. It follows that Ψ_{Dt} in (63) is a better expansion parameter *for short antennas* than is Ψ_{K1}. If (63) is used as the expansion parameter instead of $\Omega-2$, the first-order formulas are

$$[R_0]_1=20\beta_0^2h^2 \text{ ohms}, \quad (64a)$$
$$[X_0]_1=[X_0]_0=-\frac{60(\Omega-2-\ln 4)}{\beta_0 h} \text{ ohms}, \quad (64b)$$

$\Omega=10$: $[R_0]_1=20\beta_0^2h^2$ ohms;

$$[X_0]_1=-\frac{60\times 6.61}{\beta_0 h} \text{ ohms}. \quad (64c)$$

As compared with the accurate values in (62), the error in (64c) is 9.3 percent in the resistance and only 0.15 percent in the reactance. Thus, the parameter Ψ_{Dt} leads to a high degree of accuracy in the reactance; as a result, the zeroth- and first-order solutions for X_0 are identical and hardly differ from the correct value. On the other hand, the first-order resistance determined using Ψ_{Dt} is considerably in error, more so than the value obtained using the less accurate expansion parameter Ψ_{K1}. However, this must not be misinterpreted. Since there is no contribution to the resistance in the zeroth-order term in the impedance, the *analytically exact* value for the first-order resistance $[R_0]_1$ is that appropriate to the assumed zeroth-order (triangular) current. And this is precisely $20\beta_0^2h^2$ ohms. It follows that an improvement in the accuracy of the *first-order* resistance can be achieved only if a distribution function is used that more accurately represents the actual current than the triangular distribution. In particular, the component in phase with the driving voltage must be included in the distribution function and a complex expansion parameter must be used. The alternative is to evaluate a higher order solution. However, since the resistance of a short antenna actually is negligible compared with the reactance, which, therefore, dominates the iteration, the effect of higher order iterations on the values of G_0 and R_0 is likely to lead to very slow improvement. It is for this reason that the separation of the problems of determining B_0 and G_0 as in (8a) and (8b) is important for short antennas.

It is interesting to compare the correct value (62) and the first-order values (61c) and (64c) of the reactance with the value obtained from a well-known formula for the static capacitance C_0 of two conductors placed end to end. The formula in question is equivalent to[8]

$$X_0=-\frac{1}{\omega C_0}=-\frac{60(\Omega-2.485)}{\beta_0 h} \text{ ohms}, \quad (65a)$$

which gives for $\Omega=10$,

$$X_0=-\frac{60\times 7.515}{\beta_0 h} \text{ ohms}. \quad (65b)$$

It is seen that this is in error by almost 14 percent as compared with the correct value given in (62). The reason for this may be found in the derivation of C_0 in (65a) which approximates the cylindrical antenna by a spheroid of revolution with an axially uniform distribution of charge that is characteristic of the spheroid but *not* of the cylinder.

VIII. COMPARISON OF THE NEW ANALYSIS WITH THE RESULTS OF THE KING-MIDDLETON EXPANSION FOR THE VERY SHORT RECEIVING ANTENNA

The application of the King-Middleton method to the receiving antenna[4,9,10] is most readily carried out if the antenna is parallel to the incident electric vector. The integral equation (37) may be solved in the same manner indicated in (58) and (59) but using the zeroth-order distribution function $g(\zeta,\zeta')=(\cos\beta_0 h-\cos\beta_0 h)/$

[8] K. Küpfmüller, *Theoretische Elektrotechnik* (J. Springer Verlag, Berlin, 1932), p. 66.

[9] R. King, Cruft Laboratory Technical Report No. 20, Harvard University (October, 1947).

[10] S. H. Dike, Technical Report No. 14, Radiation Laboratory, The Johns Hopkins University (June, 1951).

$(1-\cos\beta_0 h)$. The function $\Psi_{U1}(\zeta)$ defined by (59) with this value of $g(\zeta, \zeta')$ leads to an expansion parameter Ψ_{U1} that differs somewhat from Ψ_{K1}. Its properties are described in detail by Dike.[10] In particular, for sufficiently short antennas, Ψ_{U1} approaches $\Omega-1$ instead of $\Omega-2$, the limiting value of Ψ_{K1}. This difference is a *necessary* consequence of the fact that the distributions of current in the two cases are not the same. Actually, since the difference between Ψ_{K1} and Ψ_{U1} is relatively small for $\Omega \geq 10$ and since in any converging process of iteration the results are insensitive to small changes in the expansion parameter, no large added error is involved if it is found convenient to use Ψ_{K1} for the receiving antenna as well as for the transmitting antenna.[4]

The normalized current $u_0=I_0/U$ at the center of an unloaded receiving antenna of length $2h$ parallel to a linearly polarized electric field $E=\beta_0 U$ as given by the first-order iteration developed according to the theory of King and Middleton is[4]

$$u_0=\frac{j4\pi}{\zeta_0\Psi}\left[\frac{\left(2-\frac{\Omega}{\Psi}\right)(1-\cos\beta_0 h)+\frac{1}{\Psi}[F_1(0)-F_1(\zeta)]}{\cos\beta_0 h+\frac{1}{\Psi}F_1(\zeta)}\right], \tag{66}$$

where Ψ is defined as in (59). The functions $F_1(0)$ and $F_1(\zeta)$ are defined as for the transmitting antenna.[2-4] For the short antenna with $\beta_0 h \leqq 0.5$ the leading terms are

$$F_1(0)\doteq 3\beta_0^2h^2/2; \quad F_1(\zeta)=-\beta_0^2h^2+j2\beta_0^2h^2/3. \tag{67}$$

Since it has been shown that the component of currrent in phase with E is negligible, only the quadrature component is required. From (66) and (67), this is

$$u_0\doteq ju_0'=\frac{j2\pi\beta_0^2h^2}{\zeta_0}\left[\frac{2\Psi-\Omega+3}{\Psi^2}\right]. \tag{68}$$

If Ψ is defined by (59) using the zeroth-order distribution characteristic of the unloaded receiving antenna parallel to the electric field, *viz.*,

$$g(\zeta, \zeta')=\frac{\cos\beta_0 h\zeta'-\cos\beta_0 h}{\cos\beta_0 h\zeta-\cos\beta_0 h}\doteq\left(\frac{1-\zeta'^2}{1-\zeta^2}\right), \tag{69a}$$

the result for the short antenna is

$$\Psi=\Psi_{U1}=\Omega-1. \tag{69b}$$

Alternatively, if the vector-potential difference is used with the same distribution function (69a) to define

$$\Psi_{Dr}(\zeta)=[\Psi_{U1}(\zeta)-\Psi_{U1}(1)]/(1-\zeta^2), \tag{70a}$$

a function is obtained which is more nearly constant at its value at $\zeta=0$ than is $\Psi_{U1}(\zeta)/(1-\zeta^2)$. Thus, an expansion parameter that more nearly fulfills the King-Middleton requirements for short antennas is given by

$$\Psi_{Dr}=[\Psi_{U1}(0)-\Psi_{U1}(1)]=\Omega-3. \tag{70b}$$

This is the analog of Ψ_{Dt} for the transmitting antenna. The corresponding formulas for u_0' as defined in (68) with numerical values for $\Omega=10$ are:

$$\Psi_{U1}=\Omega-1\text{:}\ [u_0']_1=\frac{2\pi\beta_0^2h^2}{\zeta_0}\frac{(\Omega+1)}{(\Omega-1)^2}=\frac{\beta_0^2h^2}{60\times 7.36}, \tag{71a}$$

$$\Psi_{Dr}=\Omega-3\text{:}\ [u_0']_1=\frac{2\pi\beta_0^2h^2}{\zeta_0(\Omega-3)}=\frac{\beta_0^2h^2}{60\times 7.0}. \tag{71b}$$

The correct value for $\Omega=10$ is given by (50), *viz.*, $u'=\beta_0^2h^2/60\times 6.90$. Evidently (71a) is in error by 6.7 percent, (71b) by only 1.4 percent. If the analogous expansion parameter for the transmitting antenna is arbitrarily used for the receiving antenna, the corresponding values of u_0' are:

$$\Psi_{K1}=\Omega-2\text{:}\ [u_0']_1=\frac{2\pi\beta_0^2h^2}{\zeta_0}\frac{(\Omega-1)}{(\Omega-2)^2}=\frac{\beta_0^2h^2}{60\times 7.11}, \tag{72a}$$

$\Psi_{Dt}=\Omega-2-\ln 4$:

$$[u_0']_1=\frac{2\pi\beta_0^2h^2}{\zeta_0}\frac{(\Omega-1-2\ln 4)}{(\Omega-2-\ln 4)^2}=\frac{\beta_0^2h^2}{60\times 7.02}. \tag{72b}$$

Here (72a) is in error by 3 percent; (72b) by 1.7 percent. Note, however, that these errors apply only for $\beta_0 h \leqq 0.5$ and not in general to the several Ψ functions for antennas of arbitrary length! Since these values differ only slightly from the correct value, it is relatively unimportant whether an expansion parameter for the transmitting or for a receiving antenna is used even in the first-order solution for the quadrature current. In higher order solutions this difference naturally decreases further. The best first-order solution for u_0 is (71b), using Ψ_{Dr} just as in the transmitting case it is Ψ_{Dt}. For longer antennas Ψ_{U1} and Ψ_{K1} are equally satisfactory.

The effective electrical half-length $\beta_0 h_e$ of the receiving antenna when oriented parallel to a linearly polarized electric field E is given by (52). By definition this effective length of the antenna is the length by which E must be multiplied to give the equivalent lumped voltage $V=2h_eE=2\beta_0 h_e U$ which satisfies the equation

$$I_0=V/Z_0, \tag{73}$$

where I_0 is the current at the center of the *unloaded receiving* antenna and Z_0 is the impedance of the same antenna *when center-driven*. Note that the equivalent lumped voltage V is arbitrarily so defined that I_0 and Z_0 refer to *physically entirely different situations*. Experimentally I_0 must be measured on an unloaded receiving antenna, Z_0 on a center-driven antenna. Thus these two quantities *are completely independent, theoretically, and experimentally.*

Using the King-Middleton expansion parameters Ψ_{K1} for the driven antenna and Ψ_{U1} for the unloaded receiving antenna as determined from the vector potentials, the first-order effective lengths are

$$[\beta_0 h_e]_1 = \tfrac{1}{2}[u_0 Z_0]_1 \doteq \tfrac{1}{2}[u_0' X_0]_1$$

$$= \tfrac{1}{2}\beta_0 h\left[\frac{(\Omega+1)(\Omega-2)}{(\Omega-1)^2\left(1+\dfrac{\ln 4}{\Omega-2}\right)}\right]; \quad (74a)$$

this gives

$$\Omega=10: \quad [\beta_0 h_e]_1 = 0.463\beta_0 h. \quad (74b)$$

Alternatively, if the preferred parameters Ψ_{Dt} and Ψ_{Dr} (which are determined from the vector-potential differences) are used, the first-order effective length is

$$[\beta_0 h_e]_1 = \tfrac{1}{2}\beta_0 h\left[\frac{\Omega-2-\ln 4}{\Omega-3}\right], \quad (75a)$$

which gives

$$\Omega=10: \quad [\beta_0 h_e]_1 = 0.472\beta_0 h. \quad (75b)$$

A comparison of (74b) and (75b) with the accurate value, $\beta_0 h_e = 0.478$, as given in (53a), shows that (75a) is in error by 3.1 percent, (75b) by less than 1.3 percent. The very small component I_0'' in phase with E has been neglected so that $I_0 \doteq jI_0'$.

The directivity or gain is defined by (55); the correct value for $\Omega=10$ is $D=1.50$ as given in (56). Using the first-order values of $\beta_0 h_e$ and R_0 which are determined from Ψ_{K1} and Ψ_{U1}, the first-order gain is:

$$\Omega=10: \quad [D]_1 = 1.315. \quad (76a)$$

Similarly, using $[\beta_0 h_e]_1$ and $[R_1]$, as determined from Ψ_{Dt} and Ψ_{Dr},

$$\Omega=10: \quad [D]_1 = 1.337. \quad (76b)$$

Of these, the first is in error by 12.3 percent, the second by 10.9 percent. The directivity is computed from

$$D = 30(u_0' X_0)^2 / R_0, \quad (77)$$

and involves the two independent quantities u_0 and $Z_0 = R_0 + jX_0$. It is significant that the first-order values of u_0', X_0, and R_0 used in the evaluation of (74a) involved errors of 6.7 percent, 3.3 percent, and 6.9 percent, respectively. It is primarily owing to the fact that the value of u_0' is *too small in the numerator* while the independently determined value of R_0 is *too large in the denominator* that the error in the directivity is almost double the error in any one of its components. The improvement in the expansion parameter for the short transmitting antenna from $\Psi_{K1} = \Omega - 2$ to $\Psi_{Dt} = \Omega - 2 - \ln 4$ and for the short receiving antenna from $\Psi_{U1} = \Omega - 1$ to $\Psi_{Dr} = \Omega - 3$ in accordance with the King-Middleton principle, reduces the errors in u_0' and X_0 (1.4 percent and 0.15 percent, respectively), in the evaluation of (74b), but the error in the resistance is increased to 9.3 percent. Since it has been shown in the exact analysis in Sec. III that the correct value of the resistance can be obtained only if account is taken in the distribution function of the very small component of current in phase with the driving voltage, which has no effect on the value of a real expansion parameter, it appears that any attempt to adjust the expansion parameters specifically so that the correct value of the gain is obtained as demanded by Dike[10] is not well founded. This procedure would simply introduce large errors in u_0' or X_0, or both, and result in a "solution" of the short-antenna problem in which the two essential quantities, impedance and effective length, would all be in error, but this error would be so distributed that the directivity or gain is exact. Obviously, the only proper method for reducing the error in the first-order gain is to reduce the error in *each contributing component*, that is, in u_0', X_0, and R_0. The use of Ψ_{Dt} and Ψ_{Dr} in place of Ψ_{K1} and Ψ_{U1} accomplishes this aim for u_0' and X_0. It has relatively little effect on R_0, which plays no dominant role in the first few terms of the iteration since it is extremely small compared with X_0. Even with a complex distribution function and a complex expansion parameter it is questionable whether a really accurate value of R_0 could be determined by one or two iterations so long as X_0 is enormous compared with R_0. In any event, the problem of determining accurate values of u_0 and Z_0 which combine to give an *accurate value for the gain has been solved* earlier in this paper.

IX. COORDINATION OF THE RESULTS OF THE NEW ANALYSIS OF THE ELECTRICALLY SHORT ANTENNA WITH THE GENERAL RESULTS OF THE KING-MIDDLETON ANALYSIS

In Secs. II through VI, the impedance, effective length, and directivity of an electrically short antenna ($\beta_0 h < 1$) are determined quite accurately by evaluating the complex currents that satisfy the integral equations for both transmitting and receiving antennas. In Secs. VII and VIII these new results are compared with corresponding quantities derived from the first-order King-Middleton solution in the limiting case of very short antennas ($\beta_0 h$)<0.5. In general, the King-Middleton first-order values for reactance and effective length are in good agreement with the exact values whereas the resistance and the gain are not. Moreover, owing to the smallness of R_0 compared with X_0 and its consequent relative unimportance in the iteration, second-order resistances for short antennas are not likely to be much more accurate than first-order ones. Since the new analysis provides essentially exact values of R_0 and X_0 for $\beta_0 h \leqq 0.5$ and quite accurate values up to about $\beta_0 h = 1$, it is now possible to supplement the second-order King-Middleton data where these are least accurate.

In Fig. 4 are plotted theoretical points for R_0 and X_0 as computed from the King-Middleton second-order formula[2-4] together with points ($\Omega=10$) calculated

TABLE I. Resistance and reactance of short antenna; $\Omega=10$.

$\beta_0 h$	X_0 (ohms) Formula (19) for short antenna	X_0 (ohms) K.M. second-order using Ψ_{K1}	X_0 (ohms) Interpolated	R_0 (ohms) Formula (31) for short antenna	R_0 (ohms) K.M. second-order using Ψ_{K1}	R_0 (ohms) Interpolated	$20\beta_0^2h^2\left(1+\frac{2}{15}\beta_0^2h^2\right)$ $(\Omega=\infty)$
0	$-\infty$	$-\infty$	$-\infty$	0	0	0	0
0.1	−3945		−3945	0.183		0.183	0.200
0.2	−1950		−1950	0.732		0.732	0.804
0.3	−1274		−1274	1.66		1.66	1.82
0.4	− 929		− 929	2.97		2.97	3.27
0.5	− 716	−704.8	− 716	4.67	4.99	4.67	5.17
0.6	− 569		− 569	6.80		6.80	7.54
0.7	− 460	−451.3	− 460	9.35	10.30	9.4	10.44
0.8	− 373.7		− 370	12.38		12.5	13.89
0.9	− 303.5	−293.8	− 294	15.82	18.3	16.2	17.95
1.0	− 238.2		− 234	19.90		21.0	22.67
1.1		−177.4	− 177.4		30.0	27.3	28.10
1.2		−127.1	− 127.1		37.8	35.4	34.33
1.3		− 79.8	− 79.8		47.4	45.6	41.41
1.4		− 34.3	− 34.3		59.1	58.2	
1.5		+ 10.3	+ 10.3		73.6	73.6	
1.6		+ 54.7	+ 54.7		91.7	91.7	
1.7		+ 99.7	+ 99.7		114.8	114.8	

from the new formulas for short antennas. It is seen that the curves of reactance, X_0, overlap in a wide range in which both formulas are good approximations, thus confirming the accuracy of the King-Middleton second-order reactances. On the other hand, the second-order resistances are seen to be considerably greater than the more accurate values from the new formula, just as predicted. Indeed, they lie quite close to the zeroth-order curve, $[R_0]_0=20\beta_0^2h^2(1+15\beta_0^2h^2)$, up to about $\beta_0h=0.9$, where this formula ceases to be accurate even for an infinitely thin antenna. It is to be expected, therefore, that a smooth curve that satisfies the new theory for $\beta_0h\leqq0.8$ and the King-Middleton second-order values for $\beta_0h\geqq1.4$ should yield an improved approximation of the actual resistance. This curve is shown in Fig. 4 for $\Omega=10$ and the set of interpolated resistances is in Table I. These are to be preferred to the King-Middleton second-order values where they differ from these.

TABLE II. Effective lengths of short antennas.

$\beta_0 h$	$\beta_0 h_e=2\pi h_e/\lambda_0$ $\Omega=10$	$\Omega=\infty$
0	0	0
0.1	0.048	0.050
0.2	0.096	0.100
0.3	0.144	0.149
0.4	0.193	0.197
0.5	0.243	0.254
0.6	0.293	0.291
0.7	0.345	0.336
0.8	0.398	0.380
0.9	0.452	0.422
1.0	0.516	0.462
1.1	0.586	0.501
1.2	0.654	0.537
1.3	0.741	0.572
1.4	0.866	0.604
1.5	1.12	0.635
1.6	1.30	0.664
1.7	1.46	0.691
1.8	1.61	0.716

Since accurate values of the effective length of a receiving antenna parallel to a linearly polarized electric field also are available from the new analysis for $\beta_0h<1.0$, it is possible to supplement and correct the values obtained from the King-Middleton formulas.[4] Since a second-order determination of the normalized current u_0 at the center of an unloaded receiving antenna is unavailable, the effective electrical length has been evaluated using the first-order normalized current in conjunction with the second-order impedance. That is,

$$\beta_0h_e=\tfrac{1}{2}[u_0]_1[Z_0]_2. \tag{78}$$

Since u_0 and Z_0 are *independent* quantities it is evident that the best over-all approximation of β_0h_e is achieved if the best available approximations of the individual factors are used. In Fig. 6 are shown values of $|h_e/\lambda_0|$ as computed from (78) using Ψ_{K1} as the expansion parameter and from the exact formula (53a) for $\beta_0h\leqq0.5$. The two sets of data are seen to be in good agreement in the sense that the accurate values for the short antenna lie on a smooth curve joining the approximate values obtained from (78). A list of values of $|\beta_0h_e|$ for $\Omega=10$ corresponding to the smooth curve in Fig. 6 constitutes Table II.

By combining the values of $|\beta_0h_e|$ from Fig. 6 or Table II with the interpolated curve of R_0 in Fig. 4 or the values in Table I the directivity or gain may be computed using the formula

$$D=120(|\beta_0h_e|^2/R_0). \tag{79}$$

The numerical values of D so obtained are in Table III and in Fig. 7. It is seen that in the region near resonance where the effective length of thicker antennas rises sharply and crosses that of thinner ones the gain experiences an even more rapid rise since it is proportional to the square of the effective length. The *detailed* structure ($\Omega=10$) involving a minimum at the be-

ginning and a maximum at the end of the sharp rise may not be accurate. In particular, the slight dip below $D=1.5$ appears questionable. Although these details may be altered by relatively small changes in $\beta_0 h_e$ or R_0, or both, this is not true of the sharp rise. Experimental results by Taylor and Morita[11] (using a method of determining the power to a line terminated in its characteristic impedance) show a minor minimum and a minor maximum near resonance, but the entire rise is smaller than predicted in Fig. 7. On the other hand, S. H. Dike and D. D. King[12] (using a method of determining the power to a matched load—which Taylor and Morita rejected as insufficiently accurate) observe no minor extremes near resonance but, otherwise, confirm the general shape of the gain curve except for $\beta_0 h$ less than about 1.2, where the method fails. It may be concluded that only a highly accurate theoretical determination of the gain or a correspondingly accurate experimental measurement can resolve definitely the *fine structure* of the gain curve in the vicinity of resonance. Neither is available at the present time. The theoretical investigation might determine $\beta_0 h_e$ and R_0 or the distribution of current and the electric field.

X. CONCLUSION AND SIGNIFICANCE

A method has been devised and carried out in detail for evaluating all essential properties of an electrically short transmitting or receiving antenna. Unlike other techniques employed heretofore, that depend on the solution of the integral equation by iteration, the new method determines with quantitative accuracy the complete distribution of current and the impedance of a transmitting antenna and the distribution of current, the effective length, and the gain of a symmetrical receiving antenna placed parallel to the electric field. In addition to making available accurate data on the behavior of electrically short antennas, a standard of comparison is provided for other methods that are less accurate but apply to antennas that are not so severely restricted in length.

TABLE III. Directivity or gain of short antenna.

	D	
$\beta_0 h$	$\Omega=10$	$\Omega=\infty$
0	1.500	1.500
0.1	1.500	1.500
0.2	1.502	1.502
0.3	1.505	1.504
0.4	1.509	1.508
0.5	1.514	1.512
0.6	1.518	1.518
0.7	1.519	1.524
0.8	1.520	1.532
0.9	1.520	1.540
1.0	1.521	1.553
1.1	1.509	1.564
1.2	1.450	1.578
1.3	1.445	1.592
1.4	1.546	1.609
1.5	2.042	1.627
1.6	2.211	1.647
1.7	2.243	1.669
1.8	2.137	1.698

[11] R. King, Cruft Laboratory Technical Report No. 132, Harvard University (June, 1951) and Proc. Inst. Radio Engrs. **40**, 113 (1952).

[12] S. H. Dike and D. D. King, Technical Report No. 12, Radiation Laboratory, The Johns Hopkins University (May, 1951).

Letters to the Editor

Phase Changes in Silver—Tin Amalgams*

J. C. Moffett, G. Ryge, and A. G. Barkow
Physics Department, Marquette University, Milwaukee, Wisconsin
(Received June 30, 1952)

EXTENSIVE metallurgical investigation on dental amalgam, combined with some x-ray analysis, has resulted in three distinct setting processes for the silver-tin-mercury alloys. Since the usual dental amalgam contains between 50–60 percent mercury, the equation for the setting process, assuming equilibrium to be obtained, is:

$$Ag_3Sn+Hg\rightarrow\beta_1+\gamma_1+\gamma_2. \quad (1)$$

Equation (1) is evident from the phase diagram of silver-tin-mercury. However, Gayler,[1] utilizing micrographic analysis and thermal curves, found that the foregoing process is not realized and suggested instead the two-step reaction:

$$\left.\begin{aligned} Ag_3Sn+Hg&\rightarrow\beta_1+\gamma_2 \\ \beta_1+\gamma_2&\rightarrow\beta_1+\gamma_2+\gamma_1 \end{aligned}\right\}. \quad (2)$$

Troiano[2] guided by Gayler's equilibrium diagram and using x-ray methods exclusively, arrived at the reaction:

$$\left.\begin{aligned} Ag_3Sn+Hg&\rightarrow\gamma_1+\delta_2+Ag_3Sn\ (\text{unattacked}) \\ \gamma_1+\delta_2+Ag_3Sn&\rightarrow\gamma_1+\gamma_2\rightarrow\gamma_1+\gamma_2+\beta_1 \end{aligned}\right\}. \quad (3)$$

The equilibrium diagram suggests that the ultimate phases formed be beta-1, gamma-2, gamma-1; both theories must agree on this point. The point of controversy is the process by which equilibrium is finally attained. Since the dimensional changes of contraction and expansion are noted within the first few days after trituration, the identification of phases present immediately after amalgamation are most important.

The use of the recording x-ray spectrometer to obtain continual diffraction patterns of specimens from the time of trituration onward was successful and yielded new and conclusive evidence on the chronological appearance of phases. It is evident from Fig. 1 that gamma-1 and gamma-2 are present immediately after trituration. It is further evident that gamma-1 and gamma-2 are formed independently of each other. On the basis of numerous traces made on samples of varying particle sizes of the Ag_3Sn and trituration time the following reaction is suggested.

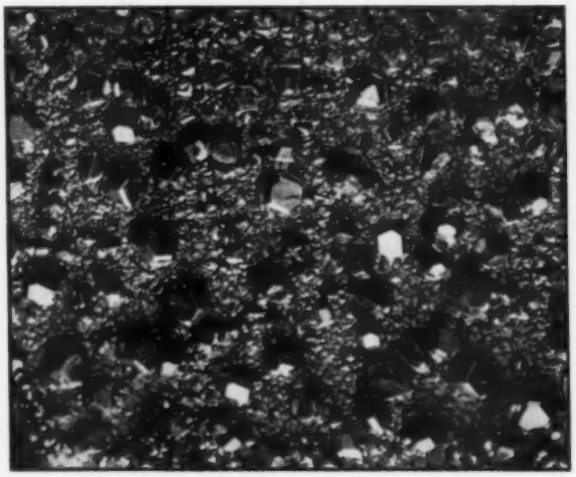


FIG. 2. Crystals of the gamma-1 phase (Ag_2Hg_3) grown on the polished surface of silver-tin (60 percent–40 percent) in contact with mercury for one month. The alloy has Ag_3Sn particles in a tin-rich matrix. Magnification 10.5×.

$$\left.\begin{aligned} Ag_3Sn+Hg&\rightarrow\gamma_1+\gamma_2+Ag_3Sn\ (\text{instantaneous}) \\ \gamma_1+\gamma_2+\gamma&\rightarrow\gamma_1+\gamma_2 \\ \gamma_1+\gamma_2&\rightarrow\gamma_1+\gamma_2+\beta_1 \end{aligned}\right\}. \quad (4)$$

There is no evidence for the last equation in (4)—it is included on the basis of the equilibrium diagram. The equations are identical with Troiano's with the exception of gamma-2 in Eq. (3). The two lines he attributed to the delta-2 phase were observed together with five other lines, but they all satisfy the gamma-2 parameters.

Efforts are being made to grow single crystals of the gamma-1 (Ag_2Hg_3), gamma-2 (Sn_7Hg) and beta-1 (Ag−Hg) phase. This is accomplished by sealing the polished face of a 60 percent Ag–40 percent Sn sample in a tray of mercury. Removal of the sample

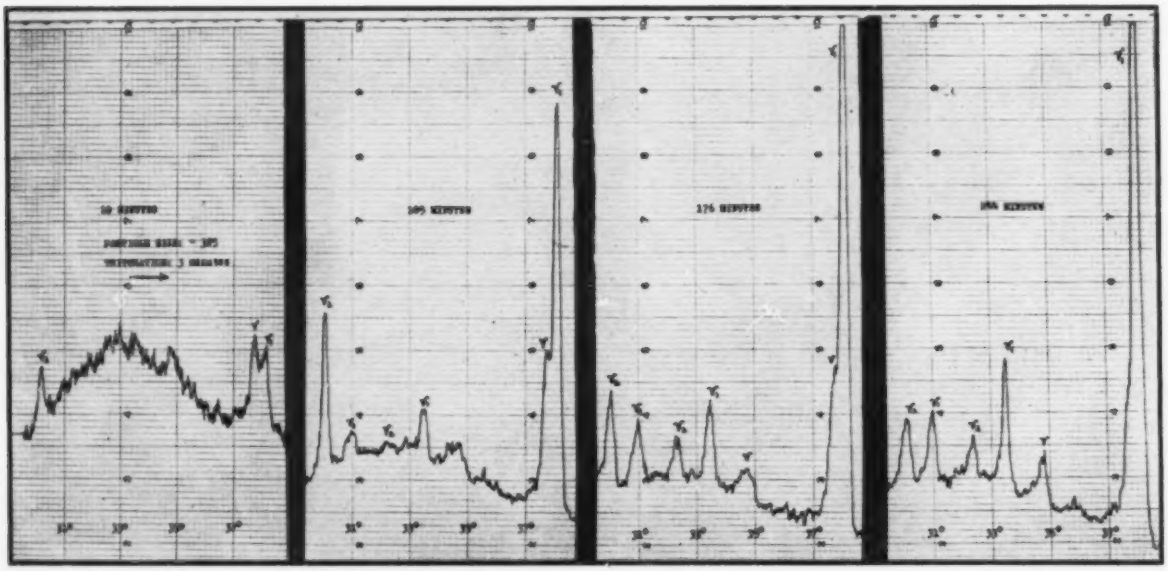


FIG. 1. X-ray spectrometer trace of 60 percent Hg and 40 percent Ag_3Sn, as a function of time after trituration. The particle size of the Ag_3Sn was less than 325 mesh. The time of preparation and adjustment of sample in the spectrometer required ten minutes. From the trace it is apparent that the gamma-1 and gamma-2 phase are present immediately. The gamma-phase gradually disappears.

FIG. 3. Crystals are grown on the polished surface of Ag_3Sn. (73.85 percent Ag, 26.15 percent Sn) in contact with mercury for one month. The gamma-1 phase (Ag_2Hg_3) and gamma-2 phase (Sn_7Hg) are suspected as being present, but crystal size is not large enough for single-crystal technique. A six-month sample is in preparation. Magnification 10.5 ×.

after one month discloses a crystal growth on the polished surface as shown in Fig. 2. The large crystals in the figure are approximately one-half millimeter cube. Investigation of crystals to date using a 5-cm radius camera and copper-K-alpha-radiation indicates an excess of gamma-1 with a lattice constant of 9.98A. The crystals appearing on the polished face of Ag_3Sn (Ag—73.85 percent, Sn—26.15 percent) after one month are shown in Fig. 3. All crystals are too small for single crystal analysis, but from the picture several structures are suspected. It is hoped that a longer growing time will result in crystals large enough to check by x-ray analysis.

* Work supported by a grant from National Institute of Dental Research.

[1] M. L. V. Gayler, J. Inst. Metals **60**, 407 (1937).

[2] A. R. Troiano, J. Inst. Metals **63**, 247 (1938).

Effect of Orientation on the Ultimate Strength of Linear High Polymers

C. C. HSIAO
The Pennsylvania State College, State College, Pennsylvania
(Received July 29, 1952)

RECENTLY a number of scientists have investigated the mechanical properties of some oriented linear high polymers. J. Bailey[1] reported some interesting results of the stretch orientation of styrene. C. C. Hsiao and J. A. Sauer[2] claimed that the effects of orientation on crazing and strength are intimately tied in with the effects of molecular structure. R. G. Cheatham and A. G. H. Dietz[3] also reported the effect of orientation on the mechanical properties of polystyrene.

The understanding of the strength of some linear high polymers has been advanced through the development of a theory which permits calculation of strength as a function of orientation in terms of strain change. Experimental results of the effect of uniaxial orientation under tension on ultimate strength are found to agree satisfactorily with theoretical calculations. Figure 1 shows the agreement of the experimental and theoretical tensile strength of polystyrene at various states of uniaxial orientation when the material is tested either in the direction of orientation or perpendicular to the direction of orientation.

The effect of uniaxial orientation on the macroscopic strength of linear high polymers is calculated by taking statistically into account the nature and arrangement of the microstructure of the material in consideration. The results of this mathematical

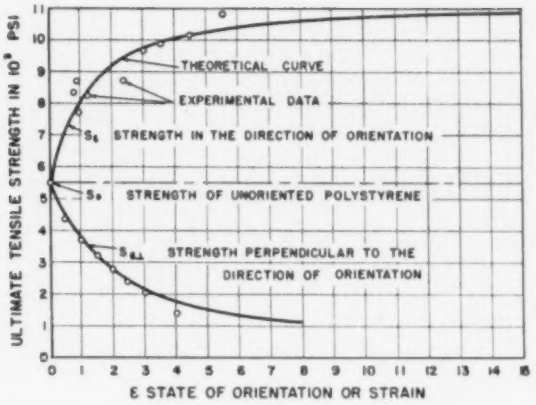

ULTIMATE TENSILE STRENGTH IN 10^3 PSI
THEORETICAL CURVE
EXPERIMENTAL DATA
S_ϵ STRENGTH IN THE DIRECTION OF ORIENTATION
S_0 STRENGTH OF UNORIENTED POLYSTYRENE
$S_{\epsilon\perp}$ STRENGTH PERPENDICULAR TO THE DIRECTION OF ORIENTATION
ε STATE OF ORIENTATION OR STRAIN

FIG. 1. Effect of uniaxial orientation on tensile strength of polystyrene.

treatment are listed for a number of different conditions of uniaxial orientation under both tension or compression.

(a) Tensile ultimate strength in the direction of uniaxial orientation produced by tension.

$$S_\epsilon = 2\frac{(1+\epsilon)^{\frac{1}{2}}}{(1+\epsilon)^{\frac{1}{2}}+1}S_0,$$

where S_ϵ is the ultimate tensile strength in the direction of orientation at the state of orientation corresponding to a strain change of ϵ in that direction, and S_0 is the strength of this material without any preferred orientation.

(b) Tensile ultimate strength perpendicular to the direction of uniaxial orientation produced by tension.

$$S_{\epsilon\perp} = 2\frac{(1+\epsilon)^{\frac{1}{2}}}{(1+\epsilon)^3-1}\cdot\frac{2}{\pi}[K-E]S_0,$$

where K and E are the values of Legendre's complete integrals of the first and second kind, respectively. If the solutions of

$$K=\frac{\pi}{2}\left[1+\left(\frac{1}{2}\right)^2k^2+\left(\frac{1\cdot3}{2\cdot4}\right)^2k^4+\left(\frac{1\cdot3\cdot5}{2\cdot4\cdot6}\right)^2k^6+\cdots\right]$$

and

$$E=\frac{\pi}{2}\left[1-\left(\frac{1}{2}\right)^2k^2-\left(\frac{1\cdot3}{2\cdot4}\right)^2k^4-\left(\frac{1\cdot3\cdot5}{2\cdot4\cdot6}\right)^2k^6-\cdots\right]$$

are used, then the relationship between the modulus k and strain ϵ is given as $k^2=1-(1+\epsilon)^{-3}$.

(c) Tensile ultimate strength in the direction of uniaxial orientation produced by compression.

$$S_{\epsilon c} = 2\frac{1}{(1+\epsilon)^{\frac{1}{2}}+1}S_0.$$

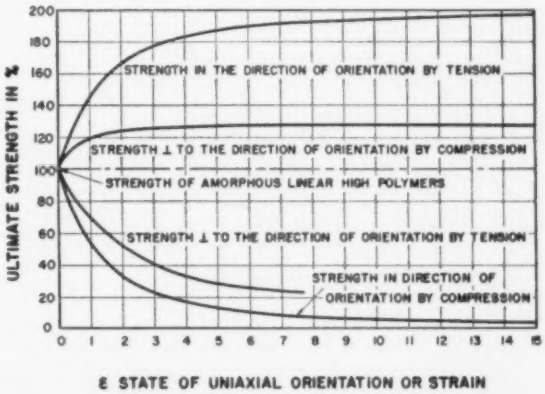

ULTIMATE STRENGTH IN %
STRENGTH IN THE DIRECTION OF ORIENTATION BY TENSION
STRENGTH ⊥ TO THE DIRECTION OF ORIENTATION BY COMPRESSION
STRENGTH OF AMORPHOUS LINEAR HIGH POLYMERS
STRENGTH ⊥ TO THE DIRECTION OF ORIENTATION BY TENSION
STRENGTH IN DIRECTION OF ORIENTATION BY COMPRESSION
ε STATE OF UNIAXIAL ORIENTATION OR STRAIN

FIG. 2. Effect of uniaxial orientation on ultimate strength of linear high polymers.

An interesting result can be obtained if the solutions for S_ϵ and $S_{\epsilon c}$ are added together for the same absolute value of ϵ. That is, the summation of these two strength quantities is constant and is equivalent to twice the strength of amorphous material.

(d) Tensile ultimate strength perpendicular to the direction of orientation produced by uniaxial compression.

$$S_{\epsilon c\perp}=\frac{4}{\pi}\left\{K-\frac{(1+\epsilon)^3}{(1+\epsilon)^3-1}[K-E]\right\}S_0.$$

These results can be plotted as shown in Fig. 2, in which the strength S_0 of amorphous linear high polymers is given in percent of 100. All the values of ultimate strength should be obtained experimentally at constant temperature and constant humidity as well as constant rate of straining for comparison.[4]

The theory applies also for cases of biaxial and triaxial orientation. Some experimental evidence of agreement has been found by the writer. However, the mathematical treatment is rather complicated. More detailed reports will be published later describing the effect of various types of orientation on the strength properties of linear high polymers.

[1] J. Bailey, India Rubber World **118**, 225 (1948).
[2] C. C. Hsiao and J. A. Sauer, J. Appl. Phys. **21**, 1071 (1950).
[3] R. G. Cheatham and A. G. H. Dietz, Trans. Am. Soc. Testing Materials (January, 1952).
[4] C. C. Hsiao and J. A. Sauer, Am. Soc. Testing Materials Bull. No. 172 (February, 1951).

On the Oscillating Cylinder Viscometer. I

ALI A. K. IBRAHIM AND A. M. KABIEL
Physics Department, Farouk Ist University, Alexandria, Egypt
(Received June 16, 1952)

CONSIDER two coaxial cylinders. The outer has an inner diameter of $2b$ cm and is made to oscillate about its axis at a given amplitude ϕ_0 and various frequencies n (in cps). The inner cylinder (l cm long with a diameter $2a$ cm) having a moment of inertia I (in g-cm²) about its major axis is suspended from a torsion fiber of restoring constant τ (in g-cm²/sec²).

When a Newtonian liquid of coefficient of viscosity η (in g/cm/sec) and density ρ (in g/cm³) is placed between the two cylinders (the annalus is completely filled). Then according to our theory[1] on the oscillating cylinder viscometer for a Newtonian liquid, we have:

$$\psi=\frac{\theta_0}{\phi_0}=\left\{\pi\cdot\frac{\rho}{d}\cdot\frac{\omega}{I\omega^2-\tau}\cdot\frac{a^3\eta}{1+d[(2\pi la^3\omega^2\rho/I\omega^2-\tau)-(3/R)]}\right\}, \quad (1)$$

where θ_0 is the amplitude of the inner cylinder, $R=(a+b)/2$, $d=(b-a)/2$, and $w=2\pi n$. Since

$$\left(\frac{d\psi}{dn}\right)_{n=0}=4\pi^2\cdot\frac{b+a}{b-a}\cdot\frac{l}{2a-b}\cdot\frac{a^3}{\tau}\cdot\eta, \quad (2)$$

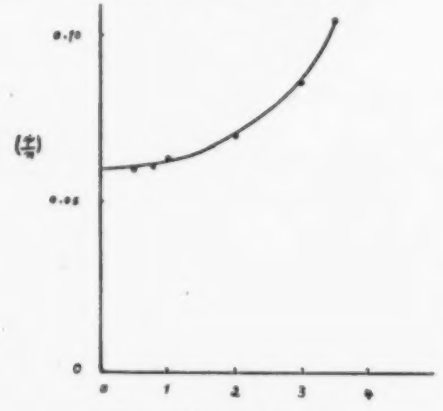


FIG. 1. $(\psi/n)_{n=0}=0.059$ sec.

therefore

$$\eta=\frac{1}{4\pi^2}\cdot\frac{b-a}{b+a}\cdot\frac{2a-b}{l}\cdot\frac{\tau}{a^3}\cdot\left(\frac{d\psi}{dn}\right)_{n=0}. \quad (3)$$

It is evident from Eq. (1) that at zero frequency of the outer cylinder the deflection θ_0 of the inner cylinder must be zero, but from Eq. (2) $(d\psi/dn)$ has a definite value even when the frequency of the outer cylinder is vanishingly small, i.e., zero.

Therefore, it seems possible that we may obtain the $(d\psi/dn)_{n=0}$ values in two ways. For reasons that will be pointed out, both methods have their advantages and disadvantages. Therefore, it has been considered advisable to use both methods, and the average of the results obtained should be closer to fact than one method alone.

First method.—First, and most simple, we may draw a tangent to the $\psi-n$ curve at the origin. The slope of this gives a measure of the quantity we need, i.e., $(d\psi/dn)_{n=0}$. This method has the advantages of being most direct but great precision cannot be claimed from this procedure.

Second method.—Second, it is possible, from the original $\psi-n$ curve, to calculate values for ψ/n. A new curve (Fig. 1) can now be

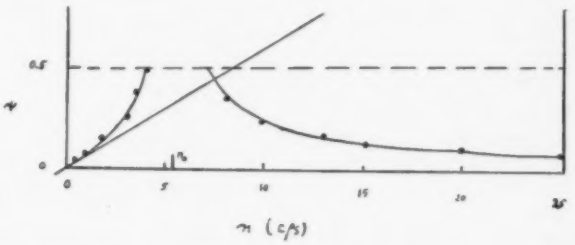


FIG. 2. $a=2.0701$ cm, $b=2.1712$ cm, $l=25.4$ cm, $\tau=2.79\ 10^8$ g-cm²/sec², $(\delta\psi/\delta n)_{n=0}=0.06$ sec.

drawn which shows the variation of ψ/n with n. Extrapolation of this curve to zero frequency (from the minimum frequency used) should lead to the quantity we need, i.e., $(\psi/n)_{n=0}$.

This method has the disadvantage of being less direct but the advantage of defining the value of $(d\psi/dn)_{n=0}$ with much more accuracy than the tangent method.

Experimental verification.—The experimental results of Oldroyd, Strawbridge, and Toms[2] (in the case of liquid Paraffin B.P. of density 0.877 g/cm³ at 25°C) may be represented in two ways (see Figs. 1 and 2). Then first, by drawing at the origin a tangent to the $\psi-n$ curve, and second, by extrapolation to zero frequency of the outer cylinder, two values of the quantity $(d\psi/dn)_{n=0}$ could be measured.

It was thus possible to use these experimental values of $(d\psi/dn)_{n=0}$ to measure the coefficient of viscosity η by substituting in Eq. (3) and comparing the experimental values of η obtained with that in standard tables.

TABLE I.

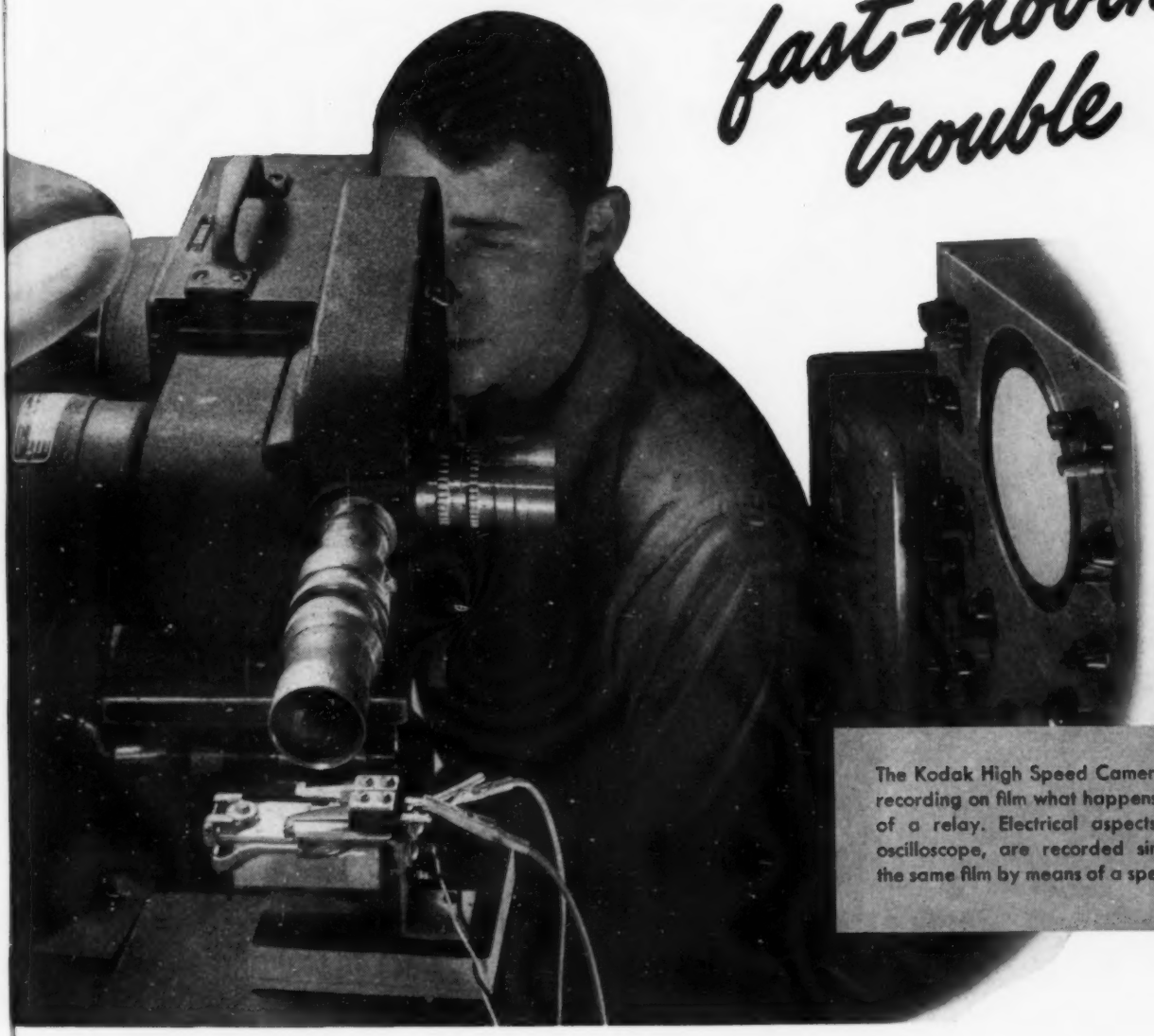
η g/cm/sec for paraffin B.P. at 25°C				
Tangent method	Extrapolation method	Mean	From standard tables	% deviation
1.7688	1.7541	1.7615	1.76	0.085

From Table I we come to the conclusion that the comparison suggests that our theory may be regarded as a satisfactory description of the facts for the nonelastic fluids.

It is a great pleasure to acknowledge the encouragement given to us by Professor Dr. M. A. Elsherbini.

[1] A. A. K. Ibrahim and A. M. Kabiel, J. Appl. Phys. **23**, 754 (1952).
[2] Oldroyd, Strawbridge, and Toms, Proc. Phys. Soc. (London) **64B**, 44 (1951).

This eye spots *fast-moving trouble*



The Kodak High Speed Camera is shown here recording on film what happens at the "break" of a relay. Electrical aspects, shown on an oscilloscope, are recorded simultaneously on the same film by means of a special attachment.

When trouble is hidden in a blur of speed too fast to see, the cause is hard to find. Here's the way to get the answer in a hurry without costly, tedious cut-and-try experimentation or theoretical analysis.

With the Kodak High Speed Camera, you can take up to 3200 clear pictures a second on 16mm film. When projected at normal speed, the film shows action slowed as much as 200 times—makes visual analysis quick and easy. And the films are available for study over and over.

This high speed "eye" is daily solving complex problems of design, production, and product performance—problems where usual methods of analysis would be slow and costly. One manufacturer projects high speed movies within two hours after they are taken—the solution to a problem is on the drawing board the same morning it is discovered. We'd be glad to send you, with our compliments, a folder showing how this company uses the Kodak High Speed Camera so effectively. Eastman Kodak Company, Industrial Photographic Division, Rochester 4, N. Y.

the Kodak *HIGH SPEED* Camera

Kodak
TRADE MARK

THE **hofman**

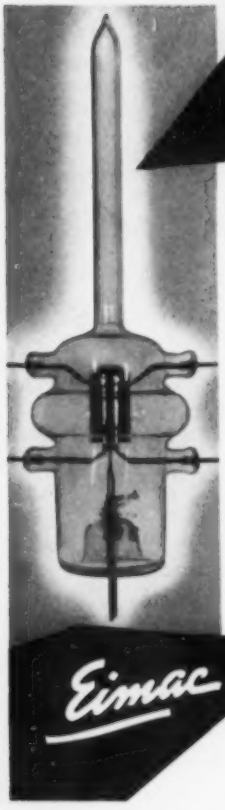
STANDARD CONTAINER

A durably constructed container for the storage and transport of liquefied gases, featuring a double wall separated by a high vacuum of 10^{-5} mm. of Hg or lower. All internal surfaces are polished to a mirror-like finish for greatest reduction of radiation losses. The strong steel outer casing is of exclusive design which eliminates damage frequently caused by rough handling. The container is suspended between heavy steel shock absorbing springs at the neck allowing free movement in any direction. Avoid costly repairs and evaporation losses — specify Hofman.

FREE! WRITE FOR BOOKLET ON HOFMAN CONTAINERS AND FLASKS

hofman LABORATORIES, INC.

212-218 WRIGHT STREET, NEWARK 5, NEW JERSEY



RUGGED— DURABLE— ADAPTABLE—

THE EIMAC 100-IG IONIZATION GAUGE

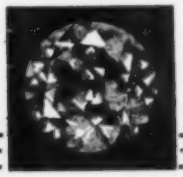
Eimac's 100-IG ionization gauge is a rugged, durable electronic pressure gauge capable of measuring pressures from 10^{-3} mm to less than 10^{-8} mm of mercury. A pure tungsten filament, gas free molybdenum grid and plate structures and a long negligible leakage path are features of the 100-IG. Its adaptability to any preferred circuit, its stability, and ability to be thoroughly out-gassed make it the most precise and practical instrument of its type.

For further information write our application engineering department.

Eimac

EITEL-McCULLOUGH, INC.

San Bruno • California

GEMS FROM  **TITANIA**

GREATER INDEX OF REFRACTION

- MORE BRILLIANT
- MORE BEAUTIFUL

THAN **DIAMONDS**

TITANIA GEMS
Completely cut and polished.
1 to 5 carats at . . . . . $12 Fed. Tax Included per carat

SPECIAL:
Limited supply of 20° Prisms
of Titania available at . . . . . $15 Each

Mounting for Ladies' and Men's Rings Available

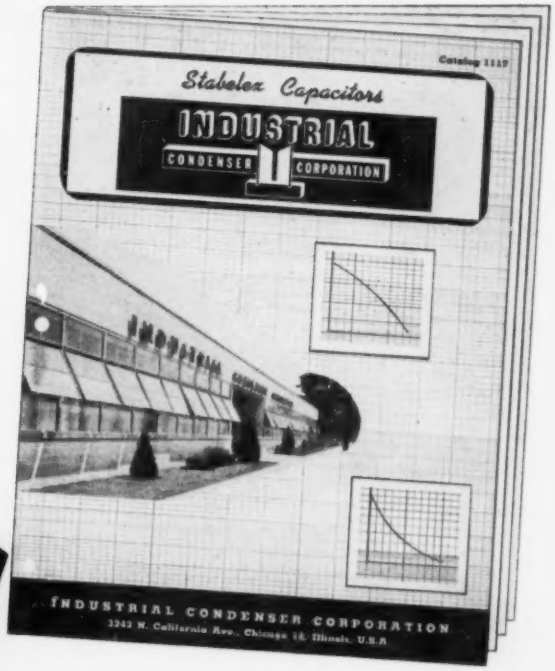
Send for FREE booklet & ring size chart

REGENT *Lapidary Co.*

Dept. JP-52, 511 East 12 St., New York 9, N. Y.

Sensational Advancements In Science & Industry

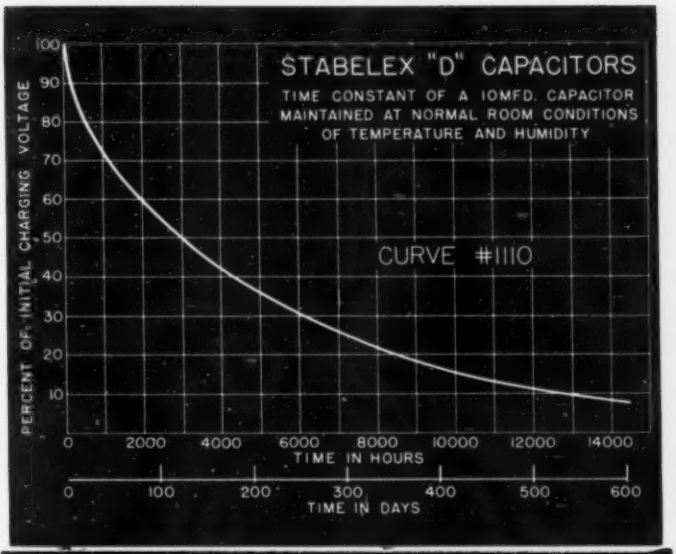
Created the Need for THE NEW *Stabelex* "D" CAPACITORS



YOUR FREE
INDUSTRIAL CONDENSER CORPORATION Stabelex "D" Capacitor Catalog may prove to be the most important new single piece of literature for you this year!

Curve #1110 is of particular interest and illustrates the long self time constant of Stabelex "D". The time constant of the 10 MFD capacitor illustrated on this curve is 200 days, or 4800 hours. This curve represents measurements on capacitors allowed to stand at normal room conditions of temperature and humidity. This, therefore, represents the time constant of these capacitors under normal conditions of operation.

Performance curves illustrating various characteristics of the Stabelex "D" Capacitor will appear in this magazine each month.



OUTSTANDING FEATURES

Insulation resistance at 20° C. after three minutes charge—900,000 megohm microfarads

Insulation resistance at 75° C.—78,000 megohm microfarads

Insulation resistance at −75° C.—In excess of 5 million megohm microfarads

Change in capacitance from 25° C. to −80° C.; +0.76%

Self time constant of 10 MFD capacitor—4800 hours

Q at 50 kilocycles—10,000

Power Factor at 1 KC—0.00025

SEND FOR CATALOG 1117 TODAY

After a long period of research, Industrial Condenser Corporation now offers to industry for the first time the first of their family of Stabelex capacitors, Stabelex "D", which has been produced for special applications for some time.

Complete information, performance curves, characteristics, and suggested applications of the various types now available will be found in this catalog.

INDUSTRIAL CONDENSER CORPORATION

3247 N. California Avenue
Chicago 18, Illinois, U.S.A.

Please send me my FREE copy of your new Catalog 1117 on Stabelex "D" Capacitors.

Name .....

Company ..... Position .....

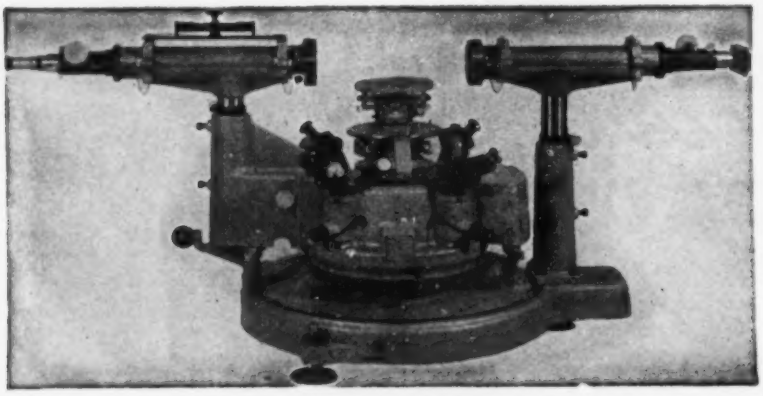
Street .....

City ..... Zone ..... State .....

MFRS. OF OIL, WAX, ELECTROLYTIC, PLASTIC CAPACITORS and RADIO INTERFERENCE FILTERS

CREEP TEST INSTRUMENTS • TELESCOPES • TOOLMAKER MICROSCOPES • COMPARATORS • OPTICAL INSTRUMENTS • SPECTROMETERS **Gaertner** REFRACTOMETERS • MICROMETER SLIDES • INTERFEROMETERS • SPECTROPHOTOMETERS • LABORATORY SUPPORTS CATHETOMETERS • CHRONOGRAPHS • OPTICAL BENCHES • SPHEROMETERS • SPECTROSCOPES • DIVIDING MACHINES PHOTOELASTICITY POLARISCOPES • MICROSCOPES • SPECTROGRAPHS • POLARIMETERS • MONOCHROMATORS

PRECISION SPECTROMETER



For highly accurate determinations of refactive indices and dispersion, measurement of prisms and wedge angles, and similar applications. Direct reading to 1 second. Polarizing and other attachments are available

A completely new instrument incorporating many adjustments contributing to the attainment of high precision and operating convenience. Cat. No. L124.

Request Bulletin 165–50

THE GAERTNER SCIENTIFIC CORPORATION

1212 WRIGHTWOOD AVE. • CHICAGO 14 • U.S.A.

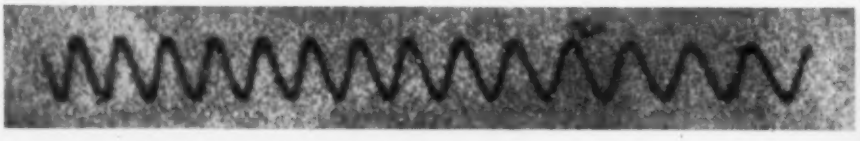
NOW! **A HIGH-SENSITIVITY, HIGH-SPEED MICROOSCILLOGRAPH**

CENTRAL RESEARCH LABORATORIES MODEL 2

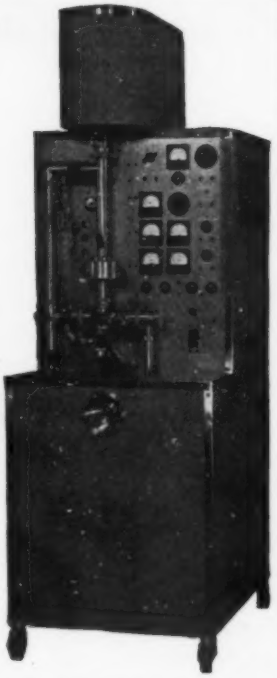
Signal-axis deflection factor only 0.2 volt per trace width!

Variation in signal-axis sensitivity less than 5 per cent from 0 to 10,000 megacycles!

Three beams, three independent deflecting systems, and direct photographic recording as in Model 1.



Unretouched single-sweep oscillogram, 3000-megacycle output of 2K41 Klystron into 50-ohm load (1.2 volts rms). Enlargement, 50 diameters.



Write us for prompt, complete information

CENTRAL RESEARCH LABORATORIES, INC., RED WING, MINN.

FUSED

QUARTZ...

for tuning forks or "hot" isotopes

Quartz is one of science's most important tools. Its many unique properties range from absolute and unchanging musical pitch, to chemical inertness to nuclear materials.

Hanovia is one of the world's major producers of clear quartz products. Hanovia's optical grades of fused quartz are universally known. But many, many other quartz laboratory tools are also available: graded joints, spherical and standard taper joints, tubing, rods, plates or special shapes to specification.

Detailed information and prices are available on all the above, upon request.

(Special Products Division)

HANOVIA CHEMICAL & MFG. CO.
DEPT., AP-10, 100 CHESTNUT ST., NEWARK 5, N. J.

Silicone

DIFFUSION PUMP FLUIDS

Dow Corning 702 and 703

Stable to Air and Moisture at 150° to 225°C

After almost 4 years' service in an electron microscope, Dow Corning 702 has yellowed slightly but is still as clear and liquid as ever.

SECOND only to mercury in chemical and thermal stability, Dow Corning Silicone Diffusion Pump Fluids withstand operating conditions which completely decompose organic oils. So resistant to oxidation and moisture that they remain clear even after distillation at atmospheric pressure, these fluids are unaffected by the accidental admission of air to a hot system. Chemically unable to break down and form carbonaceous deposits on jets, boilers or fore pumps, they remain in usable condition almost indefinitely.

Pump speeds and limiting back pressure for these silicone fluids equal those for the best organic oils, as do the ultimate vacua attainable with Dow Corning 703. For applications which do not require the lowest pressures, the economical Dow Corning 702 has a lower boiling point and operates against higher forepressure.

The rapid recovery of these fluids, plus their resistance to decomposition, means increased production and reduced maintenance costs in lens coating equipment, metallizing units, electron microscopes and vacuum tube manufacture.

Write Department CP-22 for data sheet.

DOW CORNING SILICONES

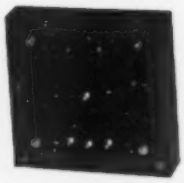
DOW CORNING CORPORATION
Midland, Michigan

NEW CONTINUOUSLY VARIABLE DELAY LINES

with

SUPERIOR

Transmission Characteristics



MINIATURE TYPE 506

- FAST RISE TIME
- EXCELLENT STABILITY
- HAIRLINE ACCURACY
- NO TIME JITTER
- SMALL SIZE
- LIGHT WEIGHT

SPECIFICATIONS

TIME DELAY: Continuously variable from 0 to 0.225 microseconds.

RISE TIME: 0.0005 $\sqrt{t}$ microsecond, where t is the amount of delay in millimicrosecond.

CHARACTERISTIC IMPEDANCE: 190 ohms nominal.

ATTENUATION: The attenuation in db per 100 millimicrosecond delay is: essentially zero below 10 mc, 0.5 at 15 mc, 1 at 20 mc, and 1.8 at 30 mc.

SIZE: 1" deep, 4" long, 4" high.

WEIGHT: 14 ounces.

TYPE 302

Type 302 distributed-parameter line offers the exclusive feature of continuously variable time delay from zero to 0.6 microsecond.

SPECIFICATIONS

TIME DELAY: Continuously variable from 0 to 0.6 microsecond.

RISE TIME: 0.00082 $\sqrt{t}$ microsecond, where t is the amount of delay in millimicrosecond.

CHARACTERISTIC IMPEDANCE: 960 ohms nominal.

ATTENUATION: The attenuation in db per 100 millimicrosecond delay is: approximately zero below 1 mc, 0.3 at 5 mc, 0.95 at 10 mc, 1.3 at 20 mc, and 1.5 at 30 mc.

WRITE FOR DATA!

Please send data on ☐ TYPE 506 ☐ TYPE 302

Name ..... Title .....

Company .....

Address .....

ADVANCE ELECTRONICS CO.

P.O. Box 394 Passaic, New Jersey

BACK ISSUES

	Single Copy
The Physical Review	
Prior to July 1929	$3.00
July 1929–December 1941	1.50
January 1942–December 1946	3.00
Thereafter	1.50
General Index, 1893–1920	4.00
Reviews of Modern Physics	2.25
Journal of the Optical Society of America	
1917–1921	2.00
Thereafter	1.00
Cumulative Index, 1917–1935	2.00
The Journal of the Acoustical Society of America	
Prior to January 1947	3.25
Thereafter	2.25
Cumulative Index, 1929–1939	4.50
Cumulative Index, 1939–1948	5.00
American Journal of Physics (*Formerly American Physics Teacher*)	
1933–1936	2.00
Thereafter	1.50
The Review of Scientific Instruments	1.00
The Journal of Chemical Physics	1.50
Journal of Applied Physics (*Formerly Physics*)	1.25

Special prices for complete volumes and complete sets.

Orders may be sent to

AMERICAN INSTITUTE OF PHYSICS

57 East 55 Street New York 22, N. Y.

Project Physicist

An outstanding opportunity for a capable, experienced senior physicist for fundamental research and development of special cathode-ray tubes.

A thorough experimental background is required with experience in electron tube and other high vacuum techniques. Experience in electron-optics and/or thin film techniques is desirable.

To work in the well-equipped Manhattan laboratory of one of the oldest electronic concerns. Top compensation for qualified man.

Write, including full details and salary required to Director of Research

Freed Radio Corporation

200 Hudson Street New York 13, N.Y.

Send for Your Free Copy

New "REVIEW" of...

- **GALVANOMETERS***
- **MICROPHOTOMETERS**
- **THERMOPILES**
- **SPECTROMETERS**
- **MONOCHROMATORS**

P. J. Kipp & Zonen, of Holland, world renowned manufacturers of these and other specialized scientific instruments, have issued the first complete catalog of their line. Your copy will be mailed as soon as we receive your request. Please ask for "Review #50-R."

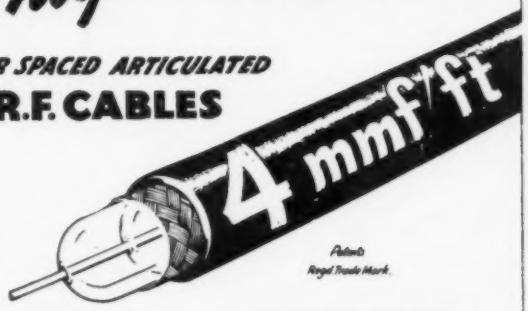
We are the exclusive United States representatives for Kipp & Zonen. Your correspondence is invited on any precision apparatus problem where we may be able to serve you.

**Immediately available in many types*

JAMES G. BIDDLE CO.
Electrical & Scientific Instruments
1316 ARCH STREET, PHILADELPHIA 7, PA.

Only with **CO-AX**

AIR SPACED ARTICULATED
R.F. CABLES



THE LOWEST EVER CAPACITANCE OR ATTENUATION

We are specially organized to give **SPOT DELIVERIES TO U.S.A.** *Cable your rush order for shipment by air freight. Settlement by your own $ check*

TRANSRADIO LTD
138A CROMWELL ROAD, LONDON S.W.7. ENGLAND
CABLES TRANSRAD LONDON

LOW ATTEN. Types	IMPED. Ohms	ATTEN. db/100ft at 100 Mcs	LOAD Cs. Kmc of 100 Mcs	QD*	
A 1	74	1.7	0.11	0.36	
A 2	74	1.3	0.24	0.44	HIGH POWER FLEXIBLE
A34	73	0.6	1.5	0.88	
LOW CAPAC. Types	**CAPAC. mmf/ft**	**IMPED. Ohms**	**ATTEN. db/100ft at 100Mcs**	**QD***	
C 1	7.3	150	2.5	0.36	
PC1	0.2	132	3.1	0.36	PHOTOCELL CABLE
C 11	6.3	173	3.2	0.36	
C 2	6.3	171	2.15	0.44	
C 22	5.5	184	2.8	0.44	
C 3	5.4	197	1.9	0.64	
C 33	4.8	220	2.4	0.64	VERY LOW CAPACITANCE
C 44	4.1	252	2.1	1.63	

Shallcross *for* PRECISION RESISTORS

Accuracy from ±1% to 0.05% as required—
Any mounting arrangement

Over 50 Standard Types Including:

- **JAN-R-93 types** including both "A" and "B" characteristics
- **High-stability types** (Tolerance 0.01%, Stability 0.003%)
- **Matched pairs and sets**
- **Miniature types**
- **Hermetically-sealed types**
- **Low-temperature co-efficient types**
- **Types with pre-determined time constants**
- **Lug-type midgets, hermetically-sealed**
- **Vertical style resistors**
- **Precision power types**
- **Multi-unit strip resistors**
- **Potted types**
- **Special card resistors**

Write for Shallcross Data Bulletin R3A

Shallcross Manufacturing Company
Collingdale, Pa.

40 KV - 0.01% STABILITY!

For deflecting a proton beam, a major nuclear research center required a power supply variable from 17 to 40 KV—D.C., with the output stable to 1 part in 10,000 over a period of weeks.

Write for Information

DESIGNED & BUILT BY BETA

Major Specifications:

Input: 117 Volts, 60 cycles 1 phase, 600 VA.

Output Voltage: 15,000-40,000 Volts, Positive Grounded. (H.V. Negative)

Output Ripple & Noise: Less than 2 volts (0.005%).

Stability: After 4 hour warmup, output voltage drift is less than 4 volts, or 0.01% over a period of 2 weeks.

Size: Control & High Voltage Tank—70"x22"x18". Feedback & Regulating Tank—60"x22"x18".

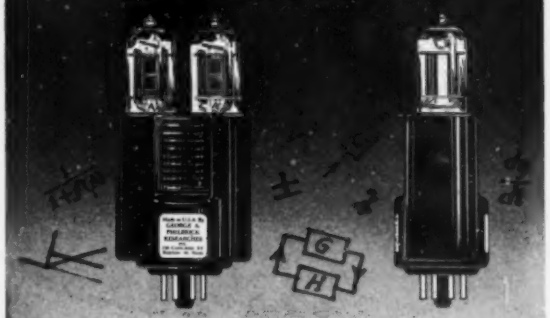
Oil Requirements: 20 gallons.

BETA Electric Corporation
333 East 103rd St. • New York, 29
ENright 9-8520

FOR DC ANALOG SYSTEMS, FAST & SLOW

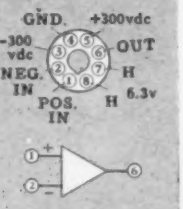
Operational Amplifier

GAP/R MODEL K2-W



SPECIFICATIONS

USAGES: Feedback Computors, dynamic synthesizers, simulators, buffer amplifiers, followers, etc., etc.
VOLTAGE GAIN: 10,000 plus.
INPUTS: Differential pair. Resistance of each is 10 megohms plus.
OUTPUTS: Impedance is 300 ohms, up to 1 milliampere and to 50 volts plus or minus.
RESPONSE: Repetitive or discrete. Aperiodic from DC up to 2 microseconds rise time under full feedback.
POWER: Total consumption 6½ watts.
TUBES: 2 12AX7. PLUG: Octal.
SEATED HEIGHT: 4 ins. WEIGHT: 2.8 oz.
UNIT PRICE: $42.40 Complete, Postpaid in USA.



George A. Philbrick Researches, Inc.
230 Congress Street, Boston 10, Massachusetts

RESEARCH ENGINEERS PHYSICISTS

The Radiation Laboratory of the Johns Hopkins University has positions in the fields of

- Electronic Circuits
- Microwaves
- Communication Theory
- Infra-Red

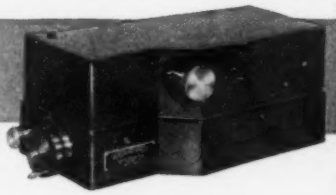
This University Laboratory offers a variety of challenging problems at both senior and junior levels. A position here means

- Favorable arrangements for advanced study in the Graduate Schools
- Faculty rank and privileges for Senior Staff
- One month paid vacation
- An air-conditioned laboratory near the University Campus

Inquiries should be addressed to:

RADIATION LABORATORY
1315 St. Paul Street
Baltimore 2, Maryland

ONE INSTRUMENT FOR THE
Ultraviolet Visible Infrared Range



FARRAND DOUBLE MONOCHROMATOR and SPECTROMETER

It is now possible to make spectral measurements in the ultraviolet, visible and infrared regions throughout the range of 0.2 to 35 microns—**WITH A SINGLE INSTRUMENT** by simply interchanging prisms.

The optical system comprises a double monochromator thereby reducing scattered radiations to a negligible magnitude. Highest spectral purity—Isolation of extremely narrow wave lengths and greatest resolution are obtained in any one chosen region by selection of prism material for optimum performance.

BULLETIN #506 ON REQUEST

PRECISION OPTICS, ELECTRONIC AND SCIENTIFIC INSTRUMENTS

FARRAND OPTICAL CO., Inc.

BRONX BLVD. and EAST 238th STREET · NEW YORK 70, N. Y.

The Roto Steppers

provide 360° clockwise and counter-clockwise rotation in 2° increments as standard. Shaft output has sufficient torque to drive many low torque mechanisms, indicators, potentiometers, selsyns, synchros, switching devices and others. Features 24 volt system, long life solenoids and combinations.

AS LISTED BY PRODUCT NO. DIA. 3.13 IN.

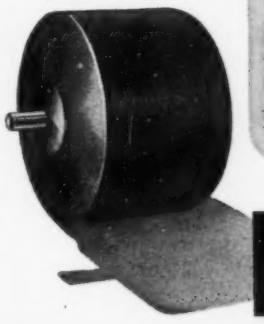
Product No.	Case Lgth.	Shaft Output	Potentiometer Output	Homing Segments	Internal Auto Cycle
0915-1	2.25	√			
0915-2	3.25	√*	√		
0915-3	4.00	√*	√	√	
0915-4	4.00	√*	√	√	√

*AVAILABLE WITHOUT SHAFT OUTPUT

Write for Engineering Data

G. M. GIANNINI & CO., INC.

Pasadena 1, Calif.



Giannini

THE LOGRINC DIGITAL GRAPH PLOTTER automatically plots one variable against another algebraically in incremental steps, in response to electrical impulses. It is ideally adapted for use as a read-out device for electronic digital computers, especially digital differential analyzers, and for use in connection with such problems as aircraft tracking and automatic data reduction.

- **plots at speeds up to 20 steps per second, in incremental steps of 1/64 of an inch.**
- **simultaneous movement on both axes in either direction.**
- **can be controlled electronically or by external or remote switches or relays.**
- **will make several carbon copies or duplicating stencil.**
- **instant manual positioning of pen and drum.**
- **takes 12" x 18" paper or continuous 12" strip.**

Mechanical simplicity . . . high reliability . . . digital accuracy . . . quick pen cartridge change . . . self-contained power supply.

Additional information supplied on request.

LOGISTICS RESEARCH COMPANY

141 South Pacific Avenue
Redondo Beach, California

ELECTRONIC PHYSICISTS
ELECTRONIC ENGINEERS

OUR STEADILY EXPANDING LABORATORY OPERATIONS
ASSURE PERMANENT POSITIONS AND UNEXCELLED
OPPORTUNITY FOR PROFESSIONAL GROWTH IN

RESEARCH & DEVELOPMENT

GUIDED MISSILES
TELEVISION
ELECTRONIC NAVIGATION
SOLID STATE PHYSICS
VACUUM TUBES
RADAR

ADDRESS INQUIRIES TO THE EMPLOYMENT DEPT.
CAPEHART FARNSWORTH CORP.
FORT WAYNE, IND.

ENGINEERS
AND
PHYSICISTS
BS-MS-Ph.D:

Responsible positions in mechanical, electrical or electronic engineering, physics or engineering physics for advanced development and design of special equipment and instruments. Prefer men with minimum of two years experience in experimental research design and development of equipment, instruments, intricate mechanisms, electronic apparatus, optical equipment, servomechanisms, control devices and allied subjects. Positions are immediate and permanent importance to our operations. Southwestern location in medium sized community. Excellent employee benefits. Reply by letter giving age, experience and other qualifications. All applications carefully considered and kept strictly confidential.

Ind. Rel. Manager, Research & Development Dept.

PHILLIPS PETROLEUM COMPANY
Bartlesville, Oklahoma

"QUARTZ RESEARCH"

Engineer or physicist experienced in preparation and testing of quartz oscillator plates wanted for investigation of electrical and some chemical properties of quartz crystals. Position offers opportunity to participate in established program for introduction of synthetic quartz crystals. Knowledge of crystallography helpful. Good salary. Submit photograph and details of education and experience to:

THE BRUSH DEVELOPMENT COMPANY
Employee Relations Dept.
3405 Perkins Ave.
Cleveland 14, Ohio

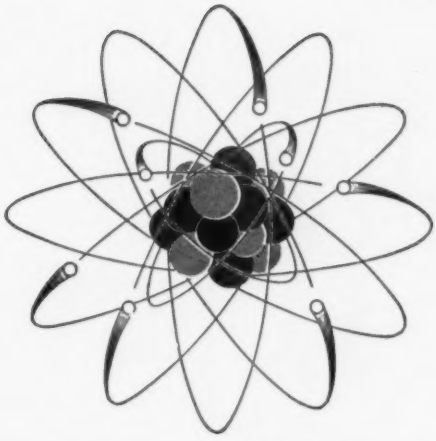
RADIOELEMENTS

For Research

high specific activity
PILE-PRODUCED ISOTOPES

naturally radioactive materials
RADIUM — RADIUM D — POLONIUM

ALPHA — BETA — GAMMA
AND NEUTRON SOURCES



Visit our booth at
NATIONAL METAL EXPOSITION
Philadelphia Convention Hall
October 20-24 1952

ATOMIC ENERGY OF CANADA LIMITED
COMMERCIAL PRODUCTS DIVISION
P.O. BOX 93 — OTTAWA, CANADA

formerly
ELDORADO
Mining and Refining (1944) Limited

EPLAB

Thermopiles

For many years the thermopile has been the accepted instrument for measuring Radiant Heat from Radiant Heaters at the American Gas Association Testing Laboratory in Cleveland, Ohio. Since 1930, when Vandaveer first described his work in this field,* an Eppley thermopile has been used for this purpose in hundreds of tests and the results have been consistent and accurate to within 1 per cent.

This is but one of many applications in the field of radiant energy measurements for which Eppley Thermopiles are ideally suited. They may be obtained with windows of different materials, and various types of black are available for receiver coatings.

All Eppley Thermopiles are supplied with a certificate of calibration, this calibration being made against a Standard Lamp from the National Bureau of Standards.

If you have a problem involving the measurement of radiant energy we invite you to write us, describing your problem in as much detail as possible. We will be glad to make recommendations and there will be no obligation.

*Vandaveer, *Industrial & Engineering Chemistry*, Vol. 22, page 596, June 1930.

BULLETIN No. 3 ON REQUEST

THE EPPLEY LABORATORY, INC.
SCIENTIFIC INSTRUMENTS
NEWPORT, RHODE ISLAND, U. S. A.

Aerophysicists, Designers, Engineers

YOUR IDEAS WILL COUNT AT NORTH AMERICAN

North American encourages advanced thinking, because they know looking ahead is the only way to maintain leadership in the aviation industry. That's why North American needs men of vision If you like hard thinking and would like to work for a company that will make the most of your ideas, you'll find real career opportunities at North American. North American offers you many extra benefits, too.

North American Extras—Salaries commensurate with ability and experience • Paid vacations • A growing organization • Complete employee service program • Cost of living bonuses • Six paid holidays a year • Finest facilities and equipment • Excellent opportunities for advancement • Group insurance including family plan • Paid sick leave • Transportation and moving allowances • Educational refund program • Low-cost group health (including family) and accident and life insurance • A company 24 years young.

Write Today—Please write us for complete information on careeer opportunities at North American. Include a summary of your education, background and experience.

IS YOUR FIELD LISTED HERE?

Airborne Electronic Equipment
Equipment Flight Tests
Precision Instruments
Automatic Controls
Propulsion Systems
Servo-Mechanisms
Airframe Studies
Radar Devices
Instrumentation
Micro Wave Techniques
Metallurgical
Electroplating
Engineering Planning

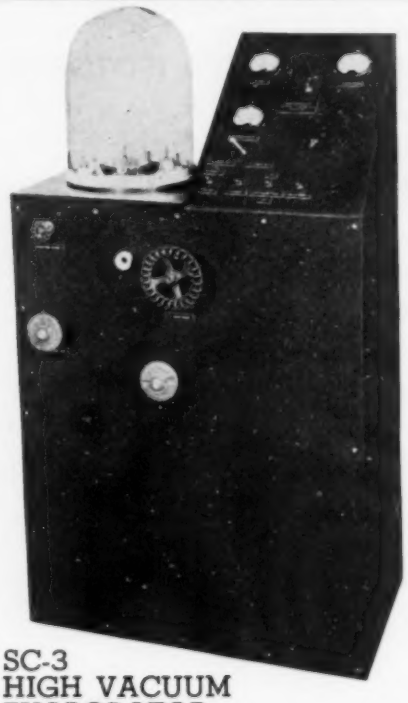
NORTH AMERICAN AVIATION, INC.

Aerophysics, Electro-Mechanical Research Division

Dept. 5, Personnel Section,
12214 Lakewood Boulevard, Downey, California

North American Has Built More Airplanes Than Any Other Company In The World

HIGH VACUUM for RESEARCH



SC-3
HIGH VACUUM
EVAPORATOR . . .

. . . combining versatility, dependability, low cost.

Among the uses of the SC-3 are . . .

- ELECTRON MICROSCOPE WORK
- MIRROR WORK
- LOW REFLECTION FILMS
- BEAM SPLITTERS
- ELECTRICAL FILMS

. . . can be used as a pumping system for other high vacuum equipment.

LARGER VACUUM UNITS (standard or custom-built) are available.

All types of vacuum coating including low reflection, front surfaced mirrors, interference filters, dichroic mirrors & crystal coatings. We are developing special coatings and will gladly quote on special filming problems.

OPTICAL FILM ENGINEERING CO.

2735-J NO. 6TH ST. PHILADELPHIA 33, PA.

DIGITAL COMPUTER RESEARCH

Your inquiry is invited concerning a position on our staff if you are experienced in the design and development of automatic data processing systems for military, industrial and commercial applications.

Logical Design
Component Development
Programming
Magnetic Recording
Circuit Design
Input & Output Devices

HUGHES

RESEARCH AND DEVELOPMENT LABORATORIES

Engineering Personnel Department

Culver City, Los Angeles County, California

Assurance is required that the relocation of the applicant will not cause the disruption of an urgent military project.

CML

Research Engineers and Senior Scientists needed to work on the University of Chicago campus in the field of Electronics. Various projects include such topics as servo systems, magnetic amplifiers, computers, and electronic and transistor circuitry.

Complete facilities and excellent associations are available in this academic atmosphere. Various research projects deal with long-range problems of the Tactical Air Command.

We offer salaries generally prevailing in industry as well as the security of a continuing position with this military laboratory. Applicants must be U. S. citizens.

For information write to:

Director of Scientific Personnel
Chicago Midway Laboratories
6040 South Greenwood Avenue
Chicago 37, Illinois

Wanted

ENGINEERS and SCIENTISTS

Unusual opportunities for outstanding and experienced men

These top positions involve preliminary and production design in advanced military aircraft and special weapons, including guided missiles.

Immediate positions include:

Boundary layer research scientists ★ Electronic project engineers ★ Electronic instrumentation engineers ★ Radar enigneers ★ Flight test engineers ★ Stress engineers ★ Weight control engineers ★ Aero- and thermodynamicists ★ Servo-mechanists ★ Power plant installation designers ★ Structural designers Electromechanical designers ★ Electrical installation designers ★ Engineering drawing checkers.

Excellent location in Southern California. Generous allowance for travel expenses.

Write today for complete information on these essential, long-term positions. Please include resume of your experience and training. Address inquiry to Director of Engineering,

NORTHROP AIRCRAFT, Inc.
1019 E. Broadway
Hawthorne (Los Angeles County) California

INDEX TO ADVERTISERS

Name	*Page*
ADVANCE ELECTRONICS CO.	xxvi
THE ARNOLD ENGINEERING COMPANY	xiii
ATOMIC INSTRUMENT COMPANY	ix
BALLANTINE LABORATORIES, INC.	ii
BEAM INSTRUMENT CORP.	iv
BELL TELEPHONE LABORATORIES	v
BETA ELECTRIC COMPANY	xxviii
JAMES G. BIDDLE COMPANY	xxvii
BOEING AIRPLANE COMPANY	xi
BRUSH DEVELOPMENT COMPANY	xxx
CAPEHART-FARNSWORTH CORP.	xxx
CENTRAL RESEARCH LABORATORIES, INC.	xxiv
CENTRAL SCIENTIFIC COMPANY	Cover 4
CHICAGO MIDWAY LABORATORIES	xxxiii
CONSOLIDATED VULTEE AIRCRAFT CORP.	x
DISTILLATION PRODUCTS INDUSTRIES (DIVISION OF EASTMAN KODAK)	iii
DOW CORNING CORPORATION	xxv
ALLEN B. DuMONT LABORATORIES	xiv
EASTMAN KODAK COMPANY	xxi
EITEL-McCULLOUGH, INC.	xxii
ELDORADO MINING & REFINING, LTD. (ATOMIC ENERGY OF CANADA, LTD.)	xxxi
THE EPPLEY LABORATORY, INC.	xxxi
FARRAND OPTICAL COMPANY	xxix
FREED RADIO CORP.	xxvi
THE GAERTNER SCIENTIFIC CORP.	xxiv
GENERAL RADIO COMPANY	Cover 3
G. M. GIANNINI & COMPANY, INC.	xxix
HANOVIA CHEMICAL & MANUFACTURING CO.	xxv
HOFMAN LABORATORIES	xxii
HUGHES RESEARCH AND DEVELOPMENT LABORATORIES	xxxiii
INDUSTRIAL CONDENSER CORP.	xxiii
THE JOHNS HOPKINS UNIVERSITY	xxviii
KINNEY MANUFACTURING COMPANY	vi
LEEDS & NORTHRUP COMPANY	i
LOGISTICS RESEARCH COMPANY	xxix
NATIONAL RESEARCH CORPORATION	xviii, xix
NORTH AMERICAN AVIATION, INC.	xxxii
NORTH AMERICAN PHILIPS COMPANY, INC.	vii
NORTHRUP AIRCRAFT, INC.	xxxiii
OPTICAL FILM ENGINEERING CO.	xxxii
G. A. PHILBRICK RESEARCHES, INC.	xxviii
PHILLIPS PETROLEUM COMPANY	xxx
RADIO CORPORATION OF AMERICA	xvii
REGENT LAPIDARY COMPANY	xxii
SHALLCROSS MANUFACTURING CO.	xxvii
SORENSEN & COMPANY, INC.	viii
TELECOMPUTING CORP.	xv
TRACERLAB, INC.	xii
TRANSRADIO, LIMITED	xxvii
W. M. WELCH SCIENTIFIC COMPANY	xx
JOHN WILEY & SONS	xvi

POSITIONS OPEN

Massachusetts Institute of Technology's Digital Computer Laboratory of the Department of Electrical Engineering has staff openings for research and development engineers and scientists, either with or without experience, for work involving logical design, theoretical analysis, component and circuit research and development (including work on transistors, ferromagnetic and ferroelectric memory cells, magnetic drums, transmission of coded information, and analog-digital conversion).

There are also staff openings for experienced people and recent graduates for work on use of high-speed digital computers to control large physical systems. This involves study of the control requirements of the whole system, reduction of these requirements to a simple pattern or sequence of control instructions, and translation of the pattern into computer code. Position requires appreciation of physical systems, ingenuity, and imagination. Prior experience with digital computers unnecessary, and training in computer principles will be provided.

Persons transferring from other fields to acquire experience in digital computers for engineering and military uses are encouraged to apply and may come on leave of absence from their permanent organizations. Position carries opportunity for academic study. Salary appropriate to candidate's experience and training. Address: Digital Computer Laboratory, MIT, 211 Massachusetts Avenue, Cambridge 39, Mass.

TECHNICAL RESEARCH ENGINEER

Experienced man with mechanical and chemical engineering background. Should have strong bent for converting research results into workable industrial equipment; should also be able to carry out essential research in this connection. Versatility and energy are prime requisites. Initial effort will be in the field of high vacuum continuous coating.

Please send resumé to:

NATIONAL RESEARCH CORPORATION

70 Memorial Drive **Cambridge 42, Mass.**

A highly successful company located in southern New England is conducting progressive, forward looking research and development in copper, titanium and other metals. We have a permanent position on our staff for a Physical Metallurgist or Physicist with a doctor's degree. While up to about 3 years' experience is desirable, we are looking more for a man of high potentialities rather than experience. If this career opportunity appeals to you, please give us full details of your technical education, experience and interests. Write Box 1052A, 57 East 55 St., New York 22, New York.

PHYSICIST

Experience in some branch of Solid State or Chemical Physics required. To formulate and conduct research on semi-conductors in conjunction with an active development program in transistors. Excellent opportunity for qualified person. Pleasant working conditions and associations. Send resume in confidence to:

The Brush Development Co.
Employee Relations Dept.
3405 Perkins Ave. Cleveland 14, Ohio

AVAILABLE

PHYSICIST—14 yrs. res. & dev. exp. in lab., theory, and management; thorough graduate training with ΦBK quality. Exp. in tubes, electron physics, ultrasonics, instrumentation, EM theory with top-notch research teams. N.Y.C. area only. Salary: $800 per month. Box 1052C, 57 East 55 Street, New York 22, New York.

Continuous Radioactivity Measurements

with the GR Counting-Rate Meter

for NUCLEAR RESEARCH ★ CHEMISTRY ★ MEDICINE & BIOLOGY
GEOLOGY ★ METALLURGY ★ AGRICULTURE ★ MINERALOGY

The G-R Type 1500-B Counting-Rate Meter, with Geiger-Mueller Counter, is a complete precision instrument for the continuous visual, aural and graphic measurement of the rate of random radiation. It is basically a laboratory instrument rather than a field survey device.

Four response speeds control meter fluctuations for varying conditions — change in rate of count occurring in a fraction of a second can be recorded or measured accurately — high input sensitivity permits use of long cable to counter tube — calibration adjustment on panel — accuracy unaffected by ±10% changes in line voltage.

WITH THIS INSTRUMENT the geologist has observed the disintegration of mineral deposits to learn the age of the earth . . . the metallurgist has compiled valuable data on case hardening, welding and alloying . . . chemists have studied photosynthesis by tracer techniques . . . biologists have determined the effects of dosage of food or of medicine on a specific organ, and have applied irradiation selectively . . . the mineralogist has tabulated the relative abundance of natural radioactive isotopes.

Crystallography, oil surveying, glass and plastic manufacturing, combustion engineering design, ore assaying and turbulence research are but few of the many fields where measurement of radioactivity is proving very valuable.

Type 1500-B
Counting-Rate Meter
$540

The Type 1500-B Counting-Rate Meter is being used to drive the Esterline-Angus 5 ma model pen recorder . . . particularly useful for obtaining a permanent graphical record of changes in rate. Visual and aural indication of radiation intensity are provided by panel meter and loud speaker.

ABRIDGED SPECIFICATIONS

Range: *full-scale values of 200, 600, 2000, 6000 and 20,000 counts per minute — minimum rate readable on meter scale is 5 counts per minute*

Pre-amplifier: *built into hand probe at end of 6-foot cable — adapter permits use of either self-quenched or internally quenched counter tubes of any design*

Accuracy: *±3% of full scale*

Response: *Four response speeds for wider range of meter damping*

Counter Circuit Voltage: *continuously variable from 400 to 2,000 volts, and available at rear of instrument — can be read from 8-position switch and calibrated dial — means provided for standardizing voltage*

Output: *trigger circuit output — recorder jack on front panel*

Accessories Supplied: *plug for connecting recorder, counter tube adapter, line cord and spare fuses (counter tubes extra)*

Type 1500-P4 Beta-Gamma-Ray Counter Tube $40
Type 1500-P5 Beta-Gamma-Ray Counter Tube $50
Type 1500-P11 Probe Mounting Stand . . . $12.50

GR®

GENERAL RADIO Company

275 Massachusetts Avenue, Cambridge 39, Massachusetts

90 West St., NEW YORK 6 · *920 S. Michigan Ave.*, CHICAGO 5 · *1000 N. Seward St.*, LOS ANGELES 38

Lifetime Accuracy

cenco MERCURIAL BAROMETERS

Above: Close-up of the conversion table in millimeters, millibars and inches.

Left: Cenco Mercurial Barometer with white lacquered metal mounting.

Accuracy . . . day in and day out . . . 'round the clock dependability . . . that's what you get in a Cenco Mercurial Barometer. Cenco features . . . patented improvements . . . precision manufacture . . . plus an effective method of filling under vacuum . . . combine to provide an instrument of unusual quality at a remarkably low price.

These instruments are built with the tube completely enclosed in hexagonal brass tubing to which the metric and English scales are attached. Patented construction of the mercury well permits fine adjustment and its transparent glass wall admits plenty of light for good visibility.

Cenco Mercurial Barometers are constructed as prescribed by the U. S. Weather Bureau and conform to correction charts established by the Bureau.

Order today for prompt delivery

No. 76890 Cenco Mercurial Barometer for altitudes 0 to 4000 ft. with scales ranging from 600 to 800 mm and 24.5 to 31.5 inches—verniers for reading to 0.1 mm or 0.005 inch. Each..... **$47.16**

No. 76891 Cenco Mercurial Barometer, same as No. 76890 but for altitudes from 1500 up to 10,000 feet. Scale ranges from 500 to 700 mm and 19.5 to 27.5 inches. Each..... **$55.15**

No. 76892 Barometer Mounting made of sheet metal for safe and convenient mounting of Cenco barometers. Finished in white lacquer to provide a light background for adjusting the zero index and vernier. With conversion table. Length, 37½ inches, width 3½ inches. Each... **$7.57**

cenco

. . . one dependable source of supply for everything you need in scientific instruments and laboratory supplies. Over 15,000 items . . . 14 branch offices and warehouses.

R.S.C. TULSA

CENTRAL SCIENTIFIC COMPANY
1700 IRVING PARK ROAD • CHICAGO 13, ILLINOIS
CHICAGO NEWARK BOSTON WASHINGTON DETROIT SAN FRANCISCO
SANTA CLARA LOS ANGELES TORONTO MONTREAL VANCOUVER OTTAWA

REFINERY SUPPLY COMPANY
624 EAST FOURTH STREET • TULSA 3, OKLAHOMA
2215 McKINNEY AVENUE • HOUSTON 3, TEXAS